

cgattcagcccggtttattttaaaaagttatgggttagatgcatgggaagaaacgcatgaa
 cgctcgattgggttgatgaaatttatcggtgtccaaccgaagatactgatcaagcaaaagca
 aaagtacaacttgcggtttgtattgatgtacgatccgaaccggttagaagacatttagaa
 agtgaagggccttttgaaacaatagggattgctgggttctttggctcgccaattcaaaaa
 5 gaagtacttgatgaacaatttgcacatccatctttaccagtgatggtagaacctgcatat
 cgtattaaagaatatgctgatcaacatgaaatgaaaatttataatcaacaacaacataca
 cttacatcttatgttttacaacttttaattaatgaaaaacaacgtgttgccaagtttgctt
 ttaccagaattaagtgggtccttttttaagtattgctgactatagctaacaacaattttccct
 aaaaaagcaaaacgtattgttcacatcgattctcacaataatggctacgtaaaccaacaggt
 10 aaattaactattcagcgtgagcaagatgcgtattcaaaactaccaatcggtttacttta
 gaagacacgattcaattttccaaaaagcattacaattaatggacttaacagatggtttt
 gcaccaacttattgtctatgctgacacggttagtgatcacataataatccctatcatgct
 tcattagagtggtgggcttgctgagtgctcgagtggttcaatgcgaattatttagca
 gtaatgtgtaatacaagaaatgttagacgtggattattgatggaaggaatcgacattcca
 15 agacatacagtggttatagctgctgaacatcaaacgtcagttgatgagttagagtattt
 tatgttccacctttaactacagaagctcaaaatgcctttgacgaacttaagcatgtgatg
 ccaaaagtatgttataaagccaatttagaacgtttggcatcggttgccaatataaataac
 actgaccataatcctaagtctgaagcgcatcgctacgctagcgattggagtgaaagttcgt
 ccagaatggggtctagcacgaaatgctgaattcattattgggaaacgtcaaatcacccaa
 20 aatagtaacttagaggagcgggcatttcttcataattatgattggacaaaggatgaagac
 ggtgagatttttaataacaattatttctgggcccagcactagtagcacaatggattaattta
 caatactacgcctcaaccgtggcacctcactattatggaagcggtagtaaaacaacgcaa
 actgtaacaagtggtgtaggtgtcatgcaaggaaatgctagtgatttaattgatggctta
 ccatggcagtcagtaattgatgaatgacaagaggcgtatcacgcacctattaggctttta
 25 attgttattcaagcgcagatgcatatattcaacggtttgttaaaacatcataatcacttt
 agacaaaagggtgatcatcaatggataagacttgccagttattgatgaaaataatagttgg
 aaagactggtag

Sequence 2010

30 VKLPYGVQQDAHEVEDALEFINRVITPLSPISTFAARNPWEGLEDASFDQVARWLKSVRD
 VDIYFNASTIHRAISNKEIDLKVFEERLDENRAHYNNRSLSDSDINTYIQRANKLMTIEE
 GYFNTKDNEKLEKWVQTNFKDKKKEDVIAQSASVFTKEGTRLIDILNAHMIKWSKLYVD
 DFQSSWTPMKREKGFYHAWQRLVKHDPLETKKQRLTLAHLPNQATEAIEYAFQELGVKEE
 HRQSYIESHLISLPGWAGIMYHRSQTQSNDAYLLTDYVAIRLSIEMVLLNDHHTLLKKS
 35 IYLQKKLEQIRYLLFNIQMNVEQWLNLSKKQAYIELGTRFSPFYFKKLWLDWEETHE
 RRLVDEIYRVPTEDTDQAKAKVQLAFCIDVRSEPFRRHLESEGGPFETIGIAGFFGLPIQK
 EVLDEQFAHPSLPVMVEPAYRIKEYADQHEMKIYNQQHTLTSMFYNFKLMKNNVLPSSL
 LPELSGPFSLIATIANITIFPKAKRIVHRFSQKWLKPTGKLTIQREQDAYSKLPIGFTL
 EEQIQFSKKALQLMDLTDDFAPLIVLCGHGSESHNNPYHASLECGACGGASSGFNAKLLA
 40 VMCNQENVRRGLLMEGIDI PRHTVFIAAEHQTSVDELEYIYVPLTTEAQNAFDELKHVM
 PKVCYKANLERLASLPNINNTDHPNNAEAHRHASDWSEVRPEWGLARNAEFIIGKRQITQ
 NSNLEGRAFLHNYDWTKDEDEILNTIISGPALVAQWINLQYYASTVAPHYYGSGSKTTQ
 TVTSGVGVMQGNASDLMYGLPWQSVMMNDKEAYHAPIRLLIVIQAPDAYIQRLLKHHNHF
 RQKVDHQWIRLASIDENNSWKDW*

Sequence 2011

Contig_0693_pos_971_603,

putative peptide of unknown function

atgaaaaagacgaaaggtatttatgaatctgaaattagtaaaagccattacacaatgggag
 50 aaagatttttaggaagaggttctttgtctgtaaaaacagatattcttcgtgatatgggt
 attgttagtttacaaggtattttgacacctgctgaatatcgctgtgtgtaaaacgaatgaa
 gggcttttaaatattaaacgtacacgttctgaactttagagtgctggtagaaggagacttg
 agtcgcattattaaagatttgactggacttaattgtgaaaagttttcatagtgatttaagt
 actattaccgggtgaacgcgtaattgatttttaagttggaagatcggttttgataaagcatta
 55 catgaagtaa

Sequence 2012

MKKTKGIYESEISKAITQWEKDFLGRGSLSVKTDILROMVIVSLQIGILTPAEYRVCKTNE
 GLLNIKRTRSELVESGEEDLSRIIKDLTGLNVKFSFHSDLSTITGERVMIFKLEDRFDKAL

HE*

Sequence 2013

Contig_0694_pos_1943_2533,

5 putative peptide of unknown function

atgatacaacgtaaagggtgaattaatactttcatggattggcaatggattacatctgttg
 tatgtatttttaataaggcatatttttcataatgacacaaacaagtgattttaaaaatggg
 atgattcaaggatttatagaagaaaatccgggggaatatgatttagcataccaacttat
 aacttgatgttaggttaggtgtgtcctcattattatacttctcattttatttaattgta
 10 tcaatagttgccgctatttttaattggcaaaaatgccaaagtatcaggaatattacttggt
 attacaggaatcattgggtctctttttaagttttattgctggtgccttatggttaatagca
 ggtatcatgcttttagtacgtaaaccacaaacacaaaatgaccaaataattctcaatat
 agtaatgacatacattcacatgttgtgccagaagaaaagaacgtgaacaacaacaatat
 aatatgaatgaaccacatattgggtcaaacatcaacatctcatcatgatcacgcattgaat
 15 gatcaaaataaacgagaaaaccataatcatgataatcaaccatacaaatag

Sequence 2014

MIQRKGELILSWIGNGLHLLYVFLIGIFFIMTQTSDFKNGMIQGFIEENPGEYDLAYQTY
 NLMGLGLGVLLIIILLILLIVSIVAAAILIGKNAKVSIGILLVITGIIGLFLSFIAGALWLIA
 20 GIMLLVRKPQTQNDQINSQYSNDIHSHVPEEKKREQQQYNMNEPHIGQTSTSHHDHALN
 DQNKRENHNHDNPYK*

Sequence 2015

Contig_0694_pos_4776_6005,

25 putative peptide of unknown function

atgcaaacagtcggaattataccttcgccagggtatagcacatcaacatgcaaaaaaata
 attccaaatgttaaacagttattgtcaaagcgtactaaacatagtcaatggaatttcgac
 atcaaagtcgatctcatgataggatctgcagaggatgtacatgaaagtgtgaaaagca
 gcacaaattaaagagggaacatcagtgaggattacgttgtttgtctgacagatttgcctagt
 30 atttcagataataaagtgggtgtcagcgactttaatagtgacaaacatgttgcaatgcta
 tcattaccgtcactaggttttattgatttgaagcgcaagctagttaaacgatgacttca
 ttgattgaacaatttatattataatcaaccgaaagacaaaaatgcgccacatccttttgta
 cgcgtgaaggctgtagaacctgacgaagacgccacatcaaaacaacgatataattaatatt
 ttatttatcataagttggattcagtttaattgggtggactgacacgagcaaatcagccttgg
 35 aaaaacatctttaattttaagaaaatcatttcagttgcctttgcaacaggaacttatgtc
 tcaatattttcaatgccatgggaattaagcgtgatttattcacgcgttcgacttatcata
 ttgatgggtgattgctatacttgggatggctggatggctattctatgcgcacatcaattgatt
 gaaagaaaaactgtcaaatctcagcgtgtatcgcataatattataattcaaccacactt
 gttacactaagtttgattacactcataaattatgtcattttatattttattgttaatcatc
 40 agtattacactctttgtccctgtggaattatttaatagttggacgagtgcccaatcacia
 ttacgttctcaaattatatgagattgatttgggttggatcatcattaggacttttagct
 ggagctatgggatcaactgttgaaaatgaagagaaaatacgtcgattacttattcttat
 agacaatatcatcgttataaagaagctggcaagaacaaaaagaacaaagaaacttctcgtg
 atgtatcacaacaaaatgtcgaacaacaaacttcaagtaaaagatgaaaataatgaacaat
 45 atgaaggtaaaaaacaaggacatagagaggagatgacgcacatgacaaatcaaaaaactgt
 ggtctagtcgtcgtccagggtgttactga

Sequence 2016

MQTVGIIPSPGIAHQHAKKIIPNVKQLLSKRTKHSQWNFDIKVDLMIGSAEDVHESVEKA
 50 AQIKEEHQWDYVVCCLDLPISISDNKVVVSDFNSDKHVAMLSLPSLGFI DLKRKL VKTMTS
 LIEQLYYNQPKDKNAPHFVRVKAVEPDEATSKQRYINILFII SWIQLIGGLTRANQPW
 KNIFNFKKIISVAFATGTYSVIFSMPELSVIYSPRLIILMVIAILGMAGWLFYAHQLI
 EKKTAKSQRVYRYIYNSTTLVTLISLITLINYVILYLLLIISITLFVPVELFNSWTSQSQ
 FTFSNYMRLIWFVSSLLAGAMGSTVENEEKIRIRITYSYRQYHRYKEAGKNKKNKLLV
 55 MYHNKMSNNKLQVKMKIMNMVKNKDIERRMTHDKSKNCGSSRRSRCY*

Sequence 2017

Contig_0696_pos_5640_5990,

is similar to (with p-value 1.0e-27)

>gp:gp|AB015981|AB015981_2 Staphylococcus aureus genes for O
rfA, MnhA, MnhB, MnhC, MnhD, MnhE, MnhF and MnhG, complete c
ds. NID: g4001723.

5 atgccagtgataagcaaaactgaataataatccgagaccatccacatataaattaaaattc
attccaatatgtggcatccaattttaagtttttcattacgttattacctgacatcgtcatt
gaaatcaatgatataaaataagacaataggaacaggtatacaaacatcct
aaatgtatacgtttaacgaaccggtataggaatggaataatgagtgcataattaacggt
aatagtaccgcaatatgtaacaaactcactatgttttctcctttaaaatttatttacga
10 cacatattatcacatcattgcgcttcttctgaaaaacatgtaactgcttga

Sequence 2018

MPVISKLNNRPSTYKLKFIPIQKFFITLLPDIVIEINDMKYINKTIGTGNTNHP
KCIRLTNRYRNGIMSANINGNSTAICNKLTMFSSFKIYLRHILYIIALLLKNHVTA*

15 Sequence 2019

Contig_0696_pos_5913_3508,

is similar to (with p-value 0.0e+00)

>gp:gp|AB015981|AB015981_2 Staphylococcus aureus genes for O
rfA, MnhA, MnhB, MnhC, MnhD, MnhE, MnhF and MnhG, complete c
20 ds. NID: g4001723.

gtgagttgttacatattgcggtactattaccgttaataatttgactcattattccattc
ctataccggttcgttaaacgtatacatttaggatggtttgtattacctgttcctattgtc
ttatttatatatttcataatcattgatttcaatgacgatgtcaggtataaacgtaataaaa
aacttaaaattggatgccacatattggaatgaattttaatttatatgtggatggtctcgga
25 ttattattcagtttgcttatcactggcataggaagtctagtgggtttatattctatcgga
tatttaagcaaatcagaacaactcggaatttttactgttatttattactatttatgggt
gcaatgttaggagttgtactctctgataattttattattttattttttgggagtta
acttctttctctagctttttatattctctttctggagagagaaaaagcttcaatttat
gggtgtcaaaaaatctttaatcattaccgtatttaggtggcttgagcatgctaggtggtatc
30 attttactttccctagctactgatacttttagtattcaggctatgatttcaaaagcaagt
gacattcaaaatagtcctttctttatcttagtaatgatactttttatgattggtgcattt
acaaaatctgcacaagtgcctttttatattttggttaccagatgctatggaggcgctacg
cczgtgagtgcataccttcattctgcaacgatggtaaaagcaggactatctaatcgca
agaatcacacctatttttgcaatatccgaaggttgggtatggacaattacacttggtggt
35 ttaatcacctatttttgggcatcactcaatgcaacaaaacaacatgacttaaaaggtatt
ttagctttctcaactgtgtctcaactagggatgattatgtctatgcttgggtattggtgct
gtaagttatcattatcaaggcgctaatagtcaactttatgttgctggatttgttgctgcc
atatcttcatatttaataatcatgccaggtttaaaggtgcactatttatgattacaggtggt
attgatcattcaactgggtacacgtgatgttaaaaaagttggcggtttacttacaatcatg
40 cctatctcattcacgcttacagttattacaacattaagtatggctggtgtgccgcctttt
aacggctttttatcaaaagagaaattcttagagtcaatgattaatgttacacatttaaat
ttaatgagtttaataactttaggtattcttttaccatcattgccattattggttagtatt
ttcacatttgtatattcaattaaatttatattgcatatattctttggttcttataaacct
gaagctctgcaaaaacaagcgcatgaatcttcaatttaattgcttatttcacctatcatt
45 ttaacatcactagttatagttatcggtttattcccaagtatattaacgcaatctattata
gagccggcatctgtagcagttagtcaaacatcaaatataactgctgagttccatttattc
catggtataaactccagcattcctatcaacaataggtatttatattattggtattttatta
ttaatttcatttagttattgggttcgtttattacaagcacatccatcagttaacgttg
aatcattggtatgacacttcaggccaacgtattccaggatattccgaaaatataacaaat
50 agttatgttacaggtttttctagaaataatttgggtgattatcttaggtattctcattgct
ttaacttttgttacagtcacagtgtaaccttcagttattgactttaaaaacgtgagtcac
ttacgcgtatttgaaggtgcaacagttattgttttactgattgcttcaactttcattata
tttgcataatcacgtttgttagcattatcatgttaagcgctgtgggttacgcta. atca
gtattattttattttctttaaagcgccgacttagcattaacacaattttagtggaatct
55 atttctacagcattattcttactatgctttttatcacctacctaatttaaatcgctacaat
gaaaaaccaacctttaaactgacaaatgctgtgatttcaattggagtggtattatcagtg
attatttttaggattaattggctatggtaatagacactttgactctattactaaattctat
caagaacatgtttttgatttagcacatggtaaaaataggtaaatgtcatcctcgtagat
ttccgtggtatggatactttattcgagtcactgtactaggtattgcaggtttaggcgta

tatacaatgattaaattacgattaaaacagaaaaatcaatcaagtgaggtgaatgaccat
gaatag

Sequence 2020

5 VSLHIAVLLPLIFALIIPFLYRFVKRIHLGWFLVPVPIVLFYIFISLISMTMSGNNVMK
NLNWMPHIGMNFNLVVDGLGLLFSLLITGIGSLVVLVSIGYLSKSEQLGNFYCYLLLFMG
AMLGVVLSDNFIILYFWELTSFSSFLISFWREKKASIYGAQKSLIITVLGGLSMLGGI
ILLSLATDTFSIQAMISKASDIQNSPFFILVMILFMIGAFTKSAQVPFYIWLDPAMEAPT
10 PVSAYLHSATMVKAGLYLIARITPIFAISEGWVWTITLVGLITLFWASLNATKQHDLCGI
LAFSTVSQGMIMSMLGIGAVSYHYQGANSQLYVAGFVAAIFHLINHATFKGALFMITGG
IDHSTGTRDVKLGLLTIPISTLTITVITLMSAGVPPFNGFLSKEKFLESMINVTHLN
LMSLNTLGILLPIITAIIGSIFTFVYSIKFILHIFGYSYKPEALPKQAHESSILMLISPII
LTSLVIVFGLFPSILTQSIIEPASVAVSQTSNITAEFHLFHGITPAFLSTIGIYIIGILL
LISFSYWVRLQLAHYPQLTLNHWDYTSQORIPGYSENITNSYVTGFSRNNLVIIILGILIA
15 LTFVTVISVPFSIDFKNVSHLRVFEGATVLFLLIASTFIIFAKSRLESIIMLSAVGYAIS
VLFIFFKAPDLATQFVVESISTALFLLCFYHLPNLNRYNEKPTFKLTNAVISIGVGLSV
IILGLIGYGNRHFSITKFYQEHVFDLAHGKMNVMNVILVDFRGMDFESSVLGIAGLV
YTMIKRLKQKNQSSEVNDHE*

20 Sequence 2021

Contig_0696_pos_3473_3087,

is similar to (with p-value 4.0e-57)

>gp:gp|AB015981|AB015981_3 Staphylococcus aureus genes for O
rfa, MnhA, MnhB, MnhC, MnhD, MnhE, MnhF and MnhG, complete c
ds. NID: g4001723.

25 gtgattatatttctcatgggtcattgtgtttggttttcattatatttagctggacattat
acacctggtggcggtttcggttggtggtttgcttttcgctagtgacattatttagttattaca
attgcatatgatgtaaaaacgatgcgaaagatttttccttttagattttaaaatcttaatt
ggatttggtttattattttgctgggtacaccatttaacaagttggttcattgtctaaaaac
30 ttttttacacatgtcacttttgacatccctttgcctttacttgaacctatgcacatgacg
acagcgatgtttttgatttcggtgttttatgtgcagttgtaggaactattatgactata
attatttcgattggagagaacgaatag

Sequence 2022

35 VIIFFMVIVFGFSLFLAGHYTPGGGFVGGLLFASALLVITIAYDVKTMRKIFPLDFKILI
GIGLLFCVGTPLTSWFMSKNFFTHVTFDIPLPLEPMHMTTAMFFDFGVLCVAVGTIMTI
IISIGENE*

Sequence 2023

40 Contig_0696_pos_3075_2740,

is similar to (with p-value 2.0e-37)

>gp:gp|AB015981|AB015981_4 Staphylococcus aureus genes for O
rfa, MnhA, MnhB, MnhC, MnhD, MnhE, MnhF and MnhG, complete c
ds. NID: g4001723.

45 atgatctttgtagtggtattctcacatctataagtgtctatctcggtttgtctaaagt
ttgatacgatcattatggggactacactactaactcatgctgcaaatattttttaatt
actatgggaggttttaagcacggaactgttccaatatttgaaaaaggaacatcaagctat
gttgaccctatcccccaagcattgattttaacagctatcggtatcgctttgctacaaca
gctttctttttagttcttgcatttagaacatataaagaactaggcactgataacgttgag
50 ctaatgaaaggagcgccagaagatgatagagagtaa

Sequence 2024

MIFVSGILTSISVYLVLKSLIRIIMGTLLTHAANLFLITMGGCLKHGTVPPIFEKGTSSY
VDPIPQALILTAIVIAFATTAFFLVLAFRYKELGTDNVELMKGAPEDDRE*

55

Sequence 2025

Contig_0696_pos_2408_1257,

is similar to (with p-value 0.0e+00)

>gp:gp|AB015981|AB015981_5 Staphylococcus aureus genes for O

rfa, MnhA, MnhB, MnhC, MnhD, MnhE, MnhF and MnhG, complete cds. **NID: g4001723.**

atgtttatggtgattggtattattgggtcatttacaacaggagatattttcaacttgttt
 gtgttctttgaagtcctttttaatgtcttcatattgtttactcgttattggtactactaaa
 atacaattacaagaacaattaagtatatatttttagtcaatgttgtttcatcgctctttcttt
 5 gtcattgggtgttgagttttatattcagttgttaggaactttaaatctcgctcatattagt
 gaaagattgtcacaactttctgtacatgacagtggttagtcaataattgtttttatttta
 tttatctttgtctttgccactaaagcaggcggttttctctatgtacgtatggctacctggt
 gcttattatgccccccagtagcgatcatcacgttctttggtgcactattgactaaagt
 10 ggtgtatacgcaattgcgagaactctaagtttattctttaataatacagtaagcttttct
 cattatgtcatccttttcttagcattacttacaattatttttgatgtataggtgcgata
 gcttactatgtatagcaagaataatcatcctttacaataattatgattgcagtaggtgtcata
 ttagttggtattgctatgatgaacgaatcaggcatgactggtgcaatatattacacacta
 catgatattgttagttaagcttcattgttcttactcattggcgctcatgtacaaaatcact
 15 aaaacgactgacttacgtcattttggtggcttgataaaaggtatcctattctaggttgg
 acattctttattgtcagcgctaagcttagcggtataccaccttttagtgggttctacgggt
 aaatctctatattgttcgagcgacctttgaaaaaggattttatctaagtggtatcatgtga
 cttttatcaagtttaatactgtgttatattcagtcatacgatttttcttaaaaggatttttc
 ggtgaagtgtgaaggatatactttatctaaaaaggtaaatgttaaatatctaacaactatc
 20 gctgttgcatctacagttattactgtaatctttggattatctgcagacacgttattccca
 atcatcaaagatggcgctgaaacgtttgtcgatccaagtcaatatattcatagtgtgtta
 ggaggtaaatag

Sequence 2026

25 MFMLIGIIGSFTTGDIFNLFVFFEVFLMSSYCLLVIGTTKIQLQETIKYILVNVVSSSFF
 VMGVAVLYSVVGTNLNLAHISERLSQLSVHDSGLVNIVFILFIFVFATKAGVFPMYVWLP
 AYYAPPVAITFFGALLTKVGVYAIARTLSLFFNNTVSFSHYVILFLALLTIIIFGCIGAI
 AYYDTKKIILYNIMIAVGVLVGIAMNESGMTGAIYYTLHDMVLKASLFLIGVYKIT
 KTTDLRHFGGLIKGYPILGWTFFFIAALSLAGIPFSGFYGKFYIVRATFEKGFYLSGIIV
 30 LLSSLIVLYSVIRIFLKGFFGEVEGYTLSKKVNVKYLTTIAVASTVITVIFGLSADTLFP
 IIKDGAETFVDPFSQYIHSVLGGK*

Sequence 2027

Contig_0696_pos_1081_776,
 35 is similar to (with p-value 3.0e-40)
 >gp:gp|AB015981|AB015981_6 Staphylococcus aureus genes for O
 rfa, MnhA, MnhB, MnhC, MnhD, MnhE, MnhF and MnhG, complete c
 ds. **NID: g4001723.**

atgcttatcattacatttttaactgagttataaaaagcaaacctttggtgtactaaaaatt
 40 attctcaaaccacgaattgagaataaacccggattctttgtgtacgagacggaattagaa
 cgtgactggcaacttggtttactttccaacttgattacgttaacacctggcacagtcgtt
 ttaggtattagtgtatgaccgtaaaaagatttatatccactcaattgatttcagtacaaag
 gaagaagagattcaaaatatcaaatcttcattagagaagggtcgttagaaaggtaggcgag
 45 aaataa

Sequence 2028

MLIITFLTELIKANFGVLKIILKPRIENKPGFFVYETELERDWQLVLLSNLITLTPGTVV
 LGISDDRKKIYIHSIDFSTKEEEIQNIKSSLEKVVVRKVGEK*

Sequence 2029

Contig_0697_pos_2646_3434,
 is similar to (with p-value 3.0e-34)
 >sp:sp|P06696|TNPA_STAAU TRANSPOSASE A (TRANSPOSON TN554). >
 pir:pir|A24584|A24584 transposition regulatory protein tnpA
 55 - Staphylococcus aureus transposon Tn554 >gp:gp|X03216|ISTN5
 54_i Staphylococcus aureus transposon Tn554. **NID: g43726.** >g
 p:gp|K02987|TRN554_1 Transposon Tn554 (from S.aureus), compl
 ete, containing transposition genes tnpA, tnpB and tnpC, and
 antibiotic resistance genes ermA and spc. **NID: g154920.**

gtgtatacttatactattaaaaattagaaatgagattattatgaaaatagtagaagtaaaa
 tctaagaatggtaccaattttatgatttttagatggtaataatgaacctatagtagatgca
 gtaagatatttgaagtatctggatagtggttaagaaaagtttaataaccaagaaaacctat
 gcctatgcactaaaaaattttttgtttacttagaaagtaaaaagatatgctataaagaa
 5 gttagttttgataaactttgttgattttataagatggatgaaaacaccttttgaatatgag
 aatgtcctctcttatcaccgaaaagaaaaagcattagtcctaagacaattaatctgact
 atgactgtagtatctaatttttatgattatctctataggagtaaaaaattagatgttaat
 ttctatgattttatgcatatggaaagtaaaactctaaaaaatataaaagtttcatgcat
 10 cacataaaataaggactatagaacggttgaaaaatatatttgaaagttaaagaaccaaagaaa
 aaaatagaagtggttaactaatgcgagggttaagaaattattagaggaagctaataatatt
 agagataaattcttaatacaattactatatgaaccggattacgtataggtgaggtatta
 tcattacgtatttgatgatattaaatttgacttttagaaacccatcgattaaaacgatacta
 ggtggtacaggaatatcaacctgttatccatcgctacgctgtcgccctcagcttagga
 cccgactaa

15

Sequence 2030

VYTYTIKIRNEIIMKIVEVKSNGTTFMILDGNNEPIVDAVRYLKYLDVSKSLNKKTY
 AYALKNFFVYLESKKICYKEVSFDNFVDFIRWMKTPFEYENVLSYHRKEKSI SPKTINLT
 MTVVSNFYDYLYRSKKLDVNFYDFMHMESKYSKKYKFSMHINKDYRTLKNILKVKEPKK
 20 KIEVLTAENVKKLLEEANNIRDKFLIQLLYETGLRIGEVLSLRIDDIKFDRNPSIKTIL
 GGTGISTCYPSPTPVGLSLGPD*

Sequence 2031

Contig_0697_pos_5460_5116,
 putative peptide of unknown function
 25 gtgacaaaaccggaggaaggtggggatgacgtcaaatacatcatgcccttatgatttgggc
 tacacacgtgctacaatggacaatacaaaagggcagcgaaactgcgaggtcaagcaaatacc
 cataaagtgtttctcagttcggattgtagtctgcaactcgactatatgaagctggatcg
 ctagtaatacgtagatcagcatgctacgggtgaatacgttccccgggtctgtacacaccgcc
 30 cgtcacaccacgagagtttgaacacccgaagccgggtggagtaaccatttggagctagcc
 gtcgaaggtgggacaaatgattgggggtgaagtcgtaacaaggtag

Sequence 2032

VTNRRKVGMTSNHHAPYDLGYTRATMDNTKGSETARSSKSHKVVLSSDCSLQLDYMKLES
 35 LVIVDQHATVNTFPLVHTARHTTRVCNTRSRWSNHLELAVEGGTNDWGEVTR*

Sequence 2033

Contig_0697_pos_1439_399,
 putative peptide of unknown function
 40 atgcaaaaagtacgttctgatataatgactcatcggtgttctcactatgatttaggagta
 aagactgctcttttggttacaaaccactcctttattaaaaaatcgaaataaagaatggcga
 aagagaattccacgttttgatattgatgtcaaggaaacctatgatataatttcaaatactat
 tcgccacagatttgggaagaaataattggatgcaagatgtattgaatctacctacaaaa
 caaatgatttttaattttggccattatcgatttactgattttaaaggacagtgggtgcaca
 45 gtatataaaggctcgatgatttttagtccgaaattatgattatcatcctgcaacatatgat
 ggtagatacttattatttcaacctaatgacgggggattatctcaaataaggaccgacttca
 agagtgactggtagaatggatggatgaacgagtatgggttagttatggcatataatttt
 atgcatcgtaaaaagcctgcaaatggattttgatgttacatgggtggcggtctaaactt
 gaaaattgcaaaaatgtaactgaagcaatcaaatttttaaagggaagtaccgcatcgtagt
 50 tcattcagttatatactaatggatagacattcgaattatgccattgtcgaagttacacct
 cgatcaatagatgtaaggtatgaacatatatgcacaaatcattttgaattgcttaccat
 gaaaatagaactatacaagagaactcaaaagacgcttaaatcgtgtaataaaataaaaca
 actccttctacaaacaaagatatcgcatcacaattatttaacgaccgcaatacgaaatc
 tatagcaacctattttaaagttgggtctgtgataaattcatacttcaactatatgaacctaat
 55 tcattaatatcatggatggcattaggtcaaaacagtcacacacctcaatcaatttttct
 aattggttaaaggaaagaaattgaatataaattactttgaaggcgaaatagatacacca
 ttaacttttgccacatactaa

Sequence 2034

5 MQKVRSDIMTHRGSHYDLGVKTALWLQTTPLLKNRNKEWRKRIPRFDIDVKETYDIFQIY
 SPQIWEEIIGMQDVLNLPKQMLNFGHYRFTDLKDSGCTVYKGRDFLVRNYDYHPATYD
 GRYLLFQPNDDGGLSQIGPISRVTGRMDGMNEYGLVMAYNFMHRRKKPANGFVCYMGRLIL
 ENCKNVTEAIKFLKEVPHRSSFSYILMDRHSNYAIVEVTPRSIDVRYEHICTNHFELLTH
 ENRNYTRESKERLNRVINKTTPSTNKDIAFKLFNDPQYEIYSNLFKSWSGTIHTSLYEPN
 SLISWMALGQNSHPTSINFNSWLKGKKNINIFYEGEIDTPLTFATY*

Sequence 2035
 Contig_0699_pos_1065_1427,
 10 is similar to (with p-value 1.0e-26)
 >gp:gp|AB003188|AB003188_3 Micrococcus luteus hexs-a, menG,
 hexs-b gene, complete cds. NID: g2982678.
 gtggcaaaagttaaacattaacaacgaaataaagaaagtagaaaagcgacttgaagaagca
 attataagttctgatcaaacattacaagaagcctcattccatttactatcttcaggggga
 15 aaaagagtttagaccgcttttgttattttaagtggtcaatttggctctaacaacaaacct
 tcagaagacacgtatcgtgtagcagtagctttagaactaatccatggctaccttagtc
 cagcatgatgtgatagataaaagtgtataaacgttagagggcgactcactatttcaaaaaaa
 tgggaccaaaagtacagctattttaacaggaaatttcttacttgctatggggctcaagcac
 tga

20 Sequence 2036
 VAKLNINNEIKKVEKRLEEAISSDQTLQEASFHLLSSGGKRVRFVILSGQFGSNNKP
 SEDTYRVAVALELIHMATLVHDDVIDKSDKRRGRILTISKWDQSTAILTGNFLLAMGLKH
 *

25 Sequence 2037
 Contig_0699_pos_2988_3362,
 is similar to (with p-value 1.0e-18)
 >sp:sp|P31114|GRC3_BACSU PROBABLE HEPTAPRENYL DIPHOSPHATE SY
 30 NTHASE COMPONENT II (EC 2.5.1.30) (HEPPP SYNTHASE) (SPORE GE
 RMINATION PROTEIN C3). >gp:gp|M80245|BACVARGNS_5 B.subtilis
 dbpA, mtr(A,B), gerC(1-3), ndk, cheR, aro(B,E,F,H), trp(A-F)
 , hisH, and tyrA genes, complete cds. NID: g143798. >gp:gp|Z
 99115|BSUB0012_214 Bacillus subtilis complete genome (sectio
 35 n 12 of 21): from 2195541 to 2409220. NID: g2634478.
 atgatcgacattataggtatgagtttccaaataatagatgatgtgctagattttact
 agttctgaaaagaaacttggttaagccggttggtagtgaccttatgaatgggtcatattaca
 ttacctgtactattagaaatgcgaaaaataagacttttaagataaaatttcacaaactt
 aatcccgacagtcctcaacatgcctttgaaacttgataacaataattagacagtccgaa
 40 agcatagaacaatacaaaacaataagtgaaaagtatttaataaagcaatcaatttaatc
 gatgaattagaggtatggtcctaataaagaactatttagaaagcttattaaaaaatggga
 agtcgaaataagtaa

Sequence 2038
 45 MIGHYIGMSFQIIDDVLDFTSSEKKLGKPVGSDLMNGHITLPVLEMRKNKTFKDKISQL
 NPDSPQHAFETCITIIRQSESEIEQSKQISEKYLKAINLIDELEDGPNKELFRKLIKMMG
 SRNK*

Sequence 2039
 50 Contig_0699_pos_4537_5598,
 is similar to (with p-value 0.0e+00)
 >sp:sp|Q59803|AROC_STAAU CHORISMATE SYNTHASE (EC 4.6.1.4) (5
 -ENOLPYRUVYLSHIKIMATE-3-PHOSPHATE PHOSPHOLYASE). >gp:gp|U319
 79|SAU31979_3 Staphylococcus aureus chorismate synthase (aro
 55 C) and nucleoside diphosphate kinase (ndk) genes, complete c
 ds, dehydroauinate synthase (aroB) and geranylgeranyl pyroph
 osphate synthetase homolog (gerCC) genes, partial cds. NID:
 g987495.
 atgtttcaaacgtcaaggggatatggtcggggcgctcgatgaaaattgaaaaagtaact

atagaaattgtgtcagggcgtcagaaatggctttacttttaggaagtccaataacttttagta
 gttactaacgacgatttcactcattggagaaaaataatggcgctcgaccaatttagtgat
 gaagaaagagaaaaatgaaacgtaccattacaaaacctagaccgggccatgctgatcct
 ataggtggcatgaaatataatcatcgtgatttaagaaatgtgcttgagcgttcacatcgtc
 5 agagaaacagcagcaagagttgctgtgggtgctgtttcaaaaattcttttagaacaatta
 gatattccacttatatagccgtgtagtcgaaattggtggtattaaagacaaagggttatat
 gatgtagatatgttcaaaaataatgtagataaaaatgatgtacgtgtaattgacgaaaat
 attgcgcaacaaatgagagataaaaatagatgaagcgaaaaagacggagattcaatcggg
 ggcgtagttcaagtaatggctgaaaacatgcctattggagtggaagtattgtacactat
 10 gaccgtaaattagatggacgcattgcacaggggtgttgtagtatcaacgccttcaaagg
 gtaagttttggtgagggatttaaaagcagctgaaaaacctggtagcgaaattcaagatgaa
 attcattataatcaagattcaggctatttttagagctacaaatcacttaggtggatttgaa
 ggaggcatgagttaatgggatgcctataattgttaattggtgcatgaagcctattcctact
 ttatataaaccactaaactcagttgatattaataactaaagaagacttcaaagctactata
 15 gaacgctcagatagttgtgcagtgcccgcagctagcgtagtagtgaacacgttgctcgt
 tttgagtttagcaaaagcagtagtcgaagagtttcaatctaaccacatggaccaactcgt
 gcacaaattaaagagcgtcgacaactcaacatagaatttttaa

Sequence 2040

20 MFKRQGGYGRGRMKIEKDTIEIVSGVRNGFTLGSPITLVVTNDDFTHWRKIMGVAPISD
 EERENMKRTITKPRPGHADLIGGMKYNHRDLRNVLERSSARETAARVAVGAVSKILLEQL
 DIHLYSRVVEIGGIKDKGLYDVMFKNNVDKNDVRVIDENIAQOMRDKIDEAKKDGSIG
 GVVQVMAENMPIGVGSYVHYDRKLDGRIAQGVVSINAFKGVSFGEFGKAAEKPGSEIQDE
 IHYNQDSGYFRATNHLGGFEGGMSNGMPIIVNGVMKPIPTLYKPLNSVDINTKEDFKATI
 25 ERSDSCAVPAASVVCEHVVA FELAKAVLEEFQSNHMDQLVAQIKERRQLNIEF*

Sequence 2041

Contig_0701_pos_1174_2742,
 is similar to (with p-value 0.0e+00)
 30 >gp:gp|Z99111|BSUB0008_149 Bacillus subtilis complete genome
 (section 8 of 21): from 1394791 to 1603020. NID: g2633699.
 >gp:gp|Z97025|BSZ97025_8 Bacillus subtilis nprE, yla[A,B,C,D
 ,E,F,G,H,I,J,K,L,M,N,O] and pycA genes. NID: g2224758.
 atggttgacgggtgctgtagttaggttgacgcataatgaaggatcaatgcctcaaaactcgt
 35 tttgttcttaaaaaagcttttagaacaacaaacttaaaaccggttgtagttgtgaataaaatt
 gataaaccagctgctagacctgagggagttgtagatgaagtattagactattcattgaa
 ttggaagcgaatgatgagcaattagacttcccagttgtttatgcttcagctgtgaatgga
 acagcaagtttagactctgaaaagcaagacgaaaatagcaatccctatacagagacgatt
 attgactatgtaccggcaccagtagataattcagatgaaccattacaattccaaattgct
 40 ttactagattataatgattatgttaggtcgtagataggcgttgacgtgtgttcagaggtaaa
 atgctgttaggtgataatgtatcactaattaaattagatggtacagttaagaactttcgt
 gtgacgaaaatatgttggttactttggtcttaaacgtgaagaaattgaagaagcacaagca
 ggagacttaatagctgtttcaggtatggaagatattaacgttggtgaaacagttacacca
 catgatcatcgtgacccattaccggtgttacgtattgatgaaccaaccctagaaatgact
 45 tttaaagtaataaactctccgtttgctggacgtgaagggtgattatgtaacagctcgacaa
 attcaagaaaagattagatcaacaacttgaaacagatgtttctttaaaagttacacctact
 gatcaaccagattcatgggttgttctggtggtggtgaactacacttgtctattcttatt
 gaaaacatgagacgtgaaggctttgaattacaggtttctaaacctcaagttattttaaga
 gaaatcgatggtgtgttaagtgaaccatttgagcgtgtacaatgtgaagtgcttctgaa
 50 aatgcggggcagtgattgagtcattaggtgcacgaaaagggtgaaatgttagatatgatg
 acgaccgacaatgggttgacgcgtttaatctttatggtacctgcacgcggtatgattggt
 tatactactgaatttatgtctatgacacgaggttatggaattattaaccatacatttgaa
 gaatttagactcgcggttaaaagctcaaatcggtggttagacgtaacggtgcattgatttct
 atggaccaaggttcaagcaacatcttatgcgattattaacttagaagatcgtgggtgtaac
 55 tttatggaaccaggtactgaagtatatgaagggtatgattgttggtgaacataaccgtgag
 aacgatttaacagtaaatattactaaagcaaacatcaaacaaacgtacgttcagctact
 aaagatcaaacacaaacgatgaatcgtcctagaattttaacattagaagaagcgttacaa
 tttatcaatgatgatgaattggtggaagtaactcctgaaagtattcgtttaagaaagaaa
 atacttaataaactctgcccgtagaaaagaagcaaaaagagttaaacaattaatgcaagac

gaacaataa

Sequence 2042

5 MVDGVVLVVDAYEGTMPQTRFVLKKALEQNLKPVVVVNKIDKPAARPEGVVDEVLDLFIE
LEANDEQLDFPVVYASAVNGTASLDSEKQDENMQSLYETIIDYVPAPVDNSDEPLQFQIA
LLDYNDYVGRIGVGRVFRGKMRVGDVSLIKLDGTVKNFRVTKIFGYFGLKREEIEEAQA
GDLIAVSGMEDINVGETVTPHHRDPLPVLRIDEPTLEMTFKVNNSPFAGREGDYVTARQ
10 IQERLDQQLTQDVSLSKVTPDQPDQSWVWVAGRGELHLSILINMRREGFELQVSKPQVILR
EIDGVLSEPFERVOCEVPSENAGAVIESLGARKGEMLDMMTTDNGLTRLI FMVPARGMIG
YTTEFMSMTRGYGIINHTFEEFRPRVKAQIGRRNGALISMDQGQATSYAIINLEDRGVN
FMEPGTEVYEGMIVGEHNRENDLTVNITKAKHQTNVRSATKDQQTQTMNRPRIITLEEALQ
FINDDELVEVTPESIRLRKKILNKSAREKEAKRVKQLMQDEQ*

Sequence 2043

15 Contig_0701_pos_3196_4209,
putative peptide of unknown function
atggaacgattttgttggtgtaaatcaaattaactatattcaaataaatccggttagaagcc
aaattttaaaccgagcgctcctaagatcatggaaaactgatcaggcagatgctcataagctt
gcttggttaggaccgacgcttaaaacacagacagcttacctatacatgagttaatattc
20 ttgaattaaagagaacgcgctccgttttcatctagaaatcgagaatgaacaaaatcgactt
aaatttcagatcccttgaattactccatcaaacattccctgggttagaaagattgttagt
agtcgatatcattcattgcactcaacatcgagaaatctttactcattcagacatgggt
cttgatatcgacaaggaggtactgattacacatatattcaattctacagataagggaatg
tcaatggataaagctacaaaatatagcacttcaattaagggtgattgctcaagaaagctat
25 cctaattgctgatagacattcctttctagtcgaaaaattacgcttacttattcaacaatta
aaacaattctattcatctcctcaacaattagatgatgccatgattcaattagcacacaaca
ctcgattattttgaaaatattcattcgatacctgggtatttgtaagctaagcacagctatg
attattggggagattgggtgatattaaagcgatttaaatcaataaaacaactcaatgctttt
gttggcattgatatacaacgatataatcaggtcatacacactgtagagataccatcaac
30 aagcgtggtataaaaaagcgagaaaacttttattttgggtgattatgaatataataaga
gggcagcatcattatgacaatcatgtcgctgattattactacaaactaagaaagcagcct
aatgagaaacctcataagactgccatcattgcttgataaatcgattattaaaaacaatt
cattatcttgaatgaatcataaattgtacgattatcaaatgtcaccacattag

35 Sequence 2044

MERFCCVNQINYIQMNPLEAKFKTSALRSWKTDQADAHKLACLGPTLKQTDSLPIHELIF
FELRERVRFHLEIENEQNRLKFQLELLHQTFPGLERLFSSRSYIIALNIAEIFTHSDMV
LDIDKEVLITHIFNSTDKGMSMDKATKYALQLRVIAQESYPNVDRHSFLVEKLRLLLIQQL
KQSIHHLKQLDDAMIQLAQQLDYFENIHSIPGIGKLSTAMIIGEIGDIKRFKSNKQLNAF
40 VGIDIKRYQSGHTHCRDTINKRGNKARKLLEFWIMNIIRGQHHYDNHVVDYKYKLRKQP
NEKPHKTAIIACINRLKTIHYLVMNHKLYDYQMSPH*

Sequence 2045

45 Contig_0701_pos_6054_6374,
putative peptide of unknown function
atgtaccggcgatgggtatctttttcaactacagctacttgcttaccactttgctttaatg
ttaacgcagcatgccaaagctgcatgaccacttcctaaaaatactacatcatattgtttca
ttaacactcatcctttcttatttttctatgagatgttttaattgtttgctctagttcttca
aacacataatttcgtttcatcactcagtcgatcaaatcttctgtgtaagcaacttgaa
50 attgcttcaaaaacagcttcttgttgttggccattgtcagttaacgtaattatcaac
tgctgctttatcagattgttga

Sequence 2046

55 MYRRWYLFQQLLAYHFALMLTQHAKLHDHFLKILHHIVSLTLILSYFSMRCFNVCSSES
NTYFVSSYSVNSCGKQLEIASKTASCCCPVSVNVIINCRSLDC*

Sequence 2047

Contig_0702_pos_13157_13459,
putative peptide of unknown function

gtgttctgtagtaggggcaaaaccaatttgataagttgtttgagttttagtttttagttat
 ctttctctctacttttttaggttgagatctacagacttgtttgccatcccttttaatt
 ttttacagttgttttcgccgttttattttggtaagaacatgtttgatatcatcaaaatt
 ttttaattttatgattacctacctgaactattttatcacctttgtgcaatccagcttcac
 5 agctggagatttcttcacaacttctcctatgacattggttggcgtaaccttggtagtatgc
 taa

Sequence 2048

VFCSRGKTNLISCLSFSFSYLSLYFFRLEIYRLVLAIPFNFYSCFRRFILVKNMFDIKI
 10 FNFMITYLNIFYFITFVQSSFISWRFLHNFSYDIGWRTLVC*

Sequence 2049

Contig_0702_pos_14035_12749,

is similar to (with p-value 2.0e-24)

15 >sp:sp|P56136|Y258_HELPY HYPOTHETICAL PROTEIN HP0258. >gp:gp
 |AE000545|HPAE000545_4 Helicobacter pylori section 23 of 134
 of the complete genome. NID: g2313349.
 atgagcatttttaattacgattatttcatctatcatcgattttggtgtactcgtaactgtt
 cacgaatatggacacatgttttttgcgaagcagcaggaattatgtgtcctgaatttgcg
 attggtatgggtcctaaaaatttttagtttgcgtaaaagatgaaacattatatacaattcgt
 20 ctattaccggtgggtggttatgtcaggatggctggtgatggtcttgaagaaccaccagtt
 caaccaggtatgaacgtaaaaaataaagttaaataaccaagacgaaatcacacataataatt
 ctagatgaccaacataaaattccaacaaattgaagccatagaagttaagaaatgtgatttt
 aaagatgacctatataattgaaggtatcacttcttatgatgatgaaaggcatcacttcact
 25 atagcgaaaaaggcattttttgtcgaaaatggaagccttggtcaaattgctccaagagat
 agacagttttacacataaagaaaccattgccaaagtttttaacattatttgcagggtccgta
 ttttaattttatttttagcttttagttctattttattggattagcatactaccaaggtagccca
 accaatgtcataggagaagttgtgaagaaatctccagctgatgaagctggattgcacaaa
 ggtgataaaatagttcaggttaggtaatacaaaaattaaaaattttgatgatatacaacat
 30 gttcttgaccaaataaaaacggcgaaaacaactgtaaaaattaaaagggatggccaaaac
 aagtctgtagatctccaacctaaaaaagtagagagaaaagataactaaaactaaaactcaa
 acaacttatcaaatgggtttttgccctactacagaacacagcggtttttaaccaataagc
 taagctatttataaactttttcgataaaaggtaagcttattttacagctgttggtgtagt
 ttagctagtataattacaggagaattttcatttgatatgttaaattggccctgttggtatt
 35 tatcacagtgttgattctgttggttaaattctggaattattaatttagtaggatacaccgct
 ttattaagtgttaacttaggaataatgaatttgctacattccagcgcttgatggtggt
 cgcatattatttgtagctatatgaggctatttttagaaaaccagtgaaataaaaagcgga
 acaggaattattgctgtaggcgcactttttgtggttattattatgatttttagtcacttgg
 aatgatatacaacggtattttcttataa

Sequence 2050

MSILITIISFIIIVFGVLVTVHEYGHMFFAKRAGIMCPEFAIGMGPKIFSFRKDETLTYTIR
 LLPVGGYVRMAGDGLLEPPVQPGMNVKIKLNNQDEITHIILDDQHKFQQIEAIEVKKCDF
 KDDLYIEGITSYDDERHHFTIAKKAFFVENGSLVQIAPRDRQFTHKKPLPKFLTLFAGPL
 45 FNFILALVLFILGAYYQGTPTNVIGEVVKKSPADEAGLHKGDKIVQVGNHKIKNFDDIKH
 VLDQNKTAKTIVKIKRDGQNKSVDLQPKKVERKITKTQTQTYQIGFAPTEHSVFKPIS
 YGIYNFFDKGLIFTAVVGMLASIFTGEFSFDMLNGPVGIIYHSVDSVVKSGIINLVGYTA
 LLSVNLGIMNLLPIPALDGGRIILFVLYEAI FRKPVNKAETGIIAVGALFVVIIMILVTW
 NDIQRYFL*

Sequence 2051

Contig_0702_pos_10795_6467,

is similar to (with p-value 0.0e+00)

55 >gp:gp|D86727|D86727_1 Staphylococcus aureus DNA for DNA pol
 ymerase III, complete cds. NID: g1483181.
 gtggttattattttggcaatgacaaatcgagaaaagtttaaagtgttgccgatcaaata
 aaaatatcaaatcaactagaacaagatattcttgaacaaggtgaactcactcgatagat
 gtttcaaataaaaacagaacatggactttccaaatatcactcccacattttttatctcat
 gaagattatcttctttttacacatgcaattgaagaagaatttaaagaaatagctacagta

gcaattgatttttcaattaaagataccaacaatcaagatgagtttgctttaaaatatttc
 ggacattgtattgatcaaacacgattgtcgccaaaagtgaaggtcaattgaaacaaaa
 aaactcattatgagtggaatgtttttaaagtcttagtttcaaagacattgagagaaat
 ctttttgataaggcatgtaattggttagtttggttaaagcatttagacagtggtgctttgaa
 5 attgataaagtcggtttttgaaacagattcaacaaatcacgatgatgacctgcatcgta
 gaagcacatattcaacaagaagatgaacaaagtgcagagaagcaactgaaaaattagaa
 aaaatgaaagcagaaaaagcgaacacaagataataatgaaagtacagtggaaaaatgt
 cagattggaaaaaccaattcagattgaaaaatataaaaccaattgaatcaattattgaagaa
 gaattcaaagtacgtattgaaggtgttatatttgatattaacctaaaagaacttaaaagt
 10 ggacgtcatatagttgagcttaaagttactgattacacagattcacttgtattaaaaatg
 ttacaagaaaaataaagatgacttggaccactttaaggcacttagtggttggtaaatgg
 gttagagctcaaggtcgattgaagaagatacttttggttagggatcttgcatgatgatg
 tcagatattgaagaaattaaaaagacacctaaacaagataaagcagaagataagcgtgta
 gagtttcatttacatacgtctatgagtcacaaatggatggatttcctaatattagtgcata
 15 gttgaacaagctgctaaatggggcaccaagcttagcagtaacagatcacacgtagta
 caagcttttctctgatgcacataatgctgccgaaaaacatggatttaagatgatttatgg
 atggaaggtatgctagtagacgatgggtgttccctatagcttataaaaccaacagaccgta
 ttaaaagatgcaacatatgtggtggttgacgtagagacaacaggtctttcctaatcaat
 gataaaatttatgaattagctgcagtaaaagtgcataacggtgaaattatagataagttt
 20 gaacgttttagtaattccacacgaaagattatctgaaaccattatcaatcttacacatac
 actgatgatattgtaactgatgctccgaaattgaagaagtgtaactgaatttaagag
 tgggttgagatgctatatattgtagctcataatgcttcatttgatagggtttattgac
 acaggatatgaaaggttaggctttggaccttctcaaacgggtgtaattgatacacttgag
 ctctcaggtacaattaataccgaatatgggaaacatgggttgaaatttcccttgccaa
 25 tatggtgctgaattaaacgcaacatcatagagcgatttatgatacagaagcaacagttat
 attttataaaaaatgggttcaacaaatgaaagaactaggtgtgaacaacctctagaatt
 aataaaaaattaaactaatgaagatgcataaaaaagagctcgtccatctcaggttacactc
 attgttcaaaatcaagaaggtcttaaaaaatttatataaaatagtttagtgcttcattagtt
 aagtattattaccgtacgccaagaattccacgttctcttttaaatgaatatcgagaagg
 30 atcttgattggtacagcttgatgaggggtgaattattcacagcagtaatgcagaaggat
 cagtcggaagtgaaaaaatagcaaaagttctatgattttatagaagttaaccgcctgcg
 ctttatcaagatttaattggatagagaattaatcagagataatgaaacgttaacacaaatt
 tacaagcgattaatagatgctggttaaagcgctaatatccagtgattgctactggtaac
 gcgcattatctatatgaacatgatgctatagccagaaaaattttaattgcatcccaacca
 35 gggaatccattaaatcggttcaacattaccagaagctcactttagaaccactgatgaaatg
 ttagatgattttcacttcttaggtgaagaaaaagcatatgaaatcggtgtaacaaataca
 aatgagctcgctaataaaaattgaaaaagtggttcctataaaaagataaaactatttacgcca
 agaattggatggggctaattgaagaaattcgtagttgagttatttcaatgcgaaaaaacta
 tatggtgaagatttaccacaaattggttatagatcgcttgaaaagggaattagatagtatt
 40 attggtaatggcttttctgttatttacctcatatctcaacgtttggtgaagaaatcgcta
 gatgatggttatttagttggatcgcggtggttcagttggttctagtctcgtagcaacaatg
 actgaaattacagaagttaatccgcttcaccacactacatttgttcacattgtaagaca
 agtgagttctttgatgatggttcggttggtatcggttcgatttaccagataaaaaatgt
 cctacttggtgtaataatgaattaaagaaggacaagatatcccttttgagacattccct
 45 ggatttaaaggagataaagttccagatattgatttgaactttagtggtgaatatcaacct
 aacgctcataattacacaaaagtattggttggtgaagataaagttatcggtgctggaaca
 ataggtactggtgctgaaaaaacagcttttggttcgtaaaaggttacttaaatgatcaa
 ggtattcacaaacgtggtgctgaaattgatcggttggttaaaggtgtacaggggtcaaa
 cgtacaactggtcaacatccctggaggaatcattgttgtagcggttacatggatatttat
 50 gattttacaccgattcaattcccagcagacgaccaaagtgcagcgtggatgacaacccat
 ttcgacttccattcaatacacgataatgtcttaaaattagatatatttaggacatgatgac
 ccaacgatgattcgatgttacaagacttatcaggaattgaccccaaaactataccagta
 gatagataaagaacaaatgcaaatatttagtggtcctgagagtttagtggttacagaagac
 gaaatattatgtaagacaggtacatttggtgtaccagaatttggtactggatttgtagct
 55 caaatgcttgaagataactaagccaacgacattctcagaattagttcaaatctcaggttta
 tctcatggtacggacgttttggttaggtaatgcacaagagtttaattcgtaagggtatgt
 gacttatctagtgtgataggctgtcggtgatgatcatgggtatctgatgatgctgga
 cttgaaccgtcaatggcttttaaacgatggaatttgtagcgtaaaaggtcggtggcttaaca
 gatgaaatgggtgaagcgatgaaggaaaataacgtgccagattgggtatttagattcttgt

cgtaaaattaaatatatgttccctaaagctcatgccgctgcttatgtactgatggctgta
 agaattgcatactttaagttacatcatccactatattattatgcagcatactttaccata
 agagcttccgattttgaccttataacaatgattaaagataaaaacgagatttcgtaataca
 gttaaagatatgtattcacgatatatggatttagggaaaaaagagaaagatgtattaact
 5 gtattagaaataatgaatgaaatggcgcatcgagggttttcgattgcaaccgattagttta
 gaaaaaagccaagcttttgacttcattgaaggggatacattgattcctccattcatt
 tcagtgccaggacttgagaaaaagcttgcacaaagaattggtgaagcgagagaagagggga
 ccatttttatccaaagaagatttaataaaaaagccggcttatctcaaaaggttattgac
 10 tatttagatgaattaggctcattgccagatttacctgacaaggcacaattgtcgatattt
 gatatgtaa

Sequence 2052

VVIIILAMTNREKFKVLADQIKISNLEQDILEQELTRIDVSNKNRTWTFQISLPHFLLSH
 EDYLLFTHAIEEFKEIATVAIDFSIKDTNNQDEFALKYFGHCIDQTRLSPKVKGQLKQK
 15 KLIMSGNVLVLSNDIERNHFDKACNGSLVKAFRQCGFEIDKVVFTDSTNHDDDIASL
 EAH:IQQEDEQSAREATEKLEKMKAEKAKQDNNESTVEKQIGKPIQIENIKPIE:IIIEE
 EFKVAIECVIFDINLKEKSGRHHVELKVTDYDLSVLKMFTRKNKDDLDHFKALSVGKW
 VRAQGRIEEDTFVRDLVMMMSDIEEIKKTPKQDKAEDKRVEFHLHTSMSQMDGIPNISAY
 VEQAAGWGHQALAVTDHNVVQAFPDAAHNAEKHGIKMIYGMGMLVDDGVPIAYKPTDRN
 20 LKDATYVVFVETGLSNQYDKIIELAAVKVHNGEIIDKFERFSNPHERLSETIINLTHI
 TDDMLTDAPEIEEVLTEFKEWVGDAIFVAHNASFDMGFIDTGYERLGFPGSTNGVIDTLE
 LSRTINTEYKGHGLNFLAKKYVELTQHHRAIYDTEATAYIFIKMVQOMKELGVNNHLEI
 NKKLTNEDAYKRARPSHVTLIVQNQEGKLNLFKIVSASLVKYYRTPRI PRSLLNEYREG
 ILIGTACDEGELFTAVMQKDQSEVEKIAKFYDFIEVQPPALYQDLMRELIRDNETLTQI
 25 YKRLIDAGKSANIPVIATGNAHYLYEHDAIARKILIASQPGNPLNRSTLPEAHFRITDEM
 LDDFHFLGEEKAYEIVVTNTNELANKIEKVPIKDKLFTPRMDGANEEIRELSYNAKKL
 YGEDLPQIVIDRLEKELDSIIGNGFSVIYLLISQRLVKKSLDDGYLVGSRGVSFVATM
 TEITEVNPLPPHYICSHCKTSEFFDDGSGVSGFDLPDKKCPTCGNELIKEGQDIPFETFL
 GFKGDKVPIDILNFSGEYQPNAHNYTKVLFGEDEKVFVFRAGTIGTVAEKTAFGFVKGYLNDQ
 30 GIHKRGAIEDRLVKGCTGVKRTTGQHPGGIIIVPDYMDIYDFTPIQFPADDQSAAWMTTH
 FDFHSIHNVKLKLDILGHDDPTMIRMLQDLSGIDPKTIPVDDKETMQIFSGPESLGVTED
 EILCKTGTGVPFEGTGFVRQMLEDTKPTTFSELVQISGLSHGTDVWLGNAQELIRQGIC
 DLSSVIGCRDDIMVYLMYAGLEPSMAFKTMEFVRKGRGLTDEMVEAMKENNVDPDWYLDSC
 RKIKYMFPAHAAAYVLMVRIAYFKVHHPLYAAAYFTIRASDFDLITMIKDKTSIRNT
 35 VKDMYSRYMDLGKKEKDVLTVLEIMNEMAHRGFRLQPI SLEKSQAFDFIIEGDTLIPFPI
 SVPGI:GENVAQRIVEAREEGPFLSKEDLNKKAGLSQKVIDYLDLGLSLPDLPKAQLSIF
 DM*

Sequence 2053

40 Contig_0702_pos_5647_4577,
 is similar to (with p-value 0.0e+00)
 >sp:sp|P32727|NUSA_BACSU N UTILIZATION SUBSTANCE PROTEIN A H
 OMOLOG (NUSA PROTEIN). >pir:pir|C36905|C36905 nusA homolog -
 Bacillus subtilis >gp:gp|Z18631|BSORF1T7A_2 B.subtilis infB
 45 -nusA operon. NID: g49314.
 atggacgaaggttcatttagagtgattgcacgtaaagaagtcgtagaagaagtggttgat
 gacagagatgaagttgatttaagtactgcttagtcaaaaatcctgcctatgaagtagga
 gatatttatgaacaagatgtaacaccgaaagacttcggacgtgtaggagctcaagcagct
 aagcaagctgtgatgcaacgacttagagacgcagaaagagaaaatttatatgatgaattt
 50 atcgataaagaagaagatattctaacaggtgtgattgaccgtgtagaccatcgctatgta
 tatgtgaatttaggaagaattgaagctgtgctgtcagaagctgaaagaagtcctaagtag
 aaatatattcctaattgaacgtatcaaggtgtacgtaaataaagttgaacagactacaaa
 ggtccacaaaatttacgtatcaagaagtcacctggattactaaaacgcttattcgaaaca
 gaagttccagaaatttatgatggtactgttattgttaaatacagtagcgcggtgaagctgga
 55 gatcgttctaaaatttagcgtgtattctgataatcctgatatagatgctgttgccgcatgt
 gtaggttctaaaggagcacgagtagaagcggttggtgaagaacttggtggcgaaaatac
 gatatcgctccaatgggatgaagatccgaaagatttgttcgtaaatgctttaagtccatca
 caagttttagaagtaattgttgatgaagagaatcaatcaactgtagttgtagttcctgat
 taccattatccttagctataggtaaaagaggcaaaacgcacgttagctgctaaatta

acaagttggaagatagatatattaatcagaatctgatgccccgagaagctggaatttatcct
gttattgaatcagaagaagttgcagatgaaattgttaattccggtgacgaagatgttgag
tttgataatgttaacttgaagagacaaacttaactagtacagaattagctgctgaaaat
gatgaagataaaaaagataaaacagaagaagataatgacacagaatcatag

5

Sequence 2054

MDEGSFRVIARKEVVEEVFDDRDEVDLSTALVKNPAYEVGDIYEQDVT PKDFGRVGAQAA
KQAVNQPLRDAEREILYDEFIDKEEDILTGVIDRVDRHYVYNLGRIEAVLSEAEASPNE
KYIPNERIKVYVNVKEQTTKGPQIYVSRSHPGLLKRLFEQEVPEIYDGTIVIVKSVAREAG
DRSKISVYSDNPDIDAVGACVSGKGARVEAVVEELGGEKIDIVQWDEDPKVFRNALSPS
QVLEIVIVDEENQSTVVVVPDYQLSLAIGKRGQNARLAAKLT SWKIDIKSES DAREAGIYP
VIESEEVADEIVNSGDEDVEFDNVNLEETNLTSTELAAENDEDDKKDKTEEDNDTES*

10

Sequence 2055

Contig_0702_pos_4266_3958,
is similar to (with p-value 3.0e-21)
>sp:sp|P32729|YLXQ_BACSU PROBABLE RIBOSOMAL PROTEIN IN NUSA-
INFB INTERGENIC REGION (ORF4). >pir:pir|E36905|E36905 hypoth
etical protein 2 (infB 5' region) - Bacillus subtilis >gp:gp
|Z18631|BSORF1T7A_4 B.subtilis infB-nusA operon. NID: g49314
. >gp:gp|Z99112|BSUB0009_132 Bacillus subtilis complete geno
me (section 9 of 21): from 1598421 to 1807200. NID: g2633902

20

atgaaaatttttaatttgcttggttagctatgagagctggtaaaatcaaaagtggcgaa
tcggtcatcttaaatgagcttaaaaagaatcaaaataaaacttgcatattagctagcgat
gcatctagtaacactctaaaacaaatgaataataatgtaatagttaccaagtgccatta
aaagtgttggtactagaaatgaattagggttagcaataggtaaaagcgatagagttaat
attggtataacagataaatggttttgcaaaaaaattgttatcaatgatagatgaatatggt
aaggagtga

30

Sequence 2056

MKIFNLLGLAMRAGKIKSGESVILNELKKNQIKLVILASDASSNTLQMNKNCNSYQVPL
KVFGTRNELGLAIGKSDRVNIGITDNGFAKKLLSMIDEYGKE*

35

Sequence 2057

Contig_0702_pos_3953_1791,
is similar to (with p-value 0.0e+00)
>sp:sp|P17889|IF2_BACSU TRANSLATION INITIATION FACTOR IF-2.
>pir:pir|A35269|A35269 translation initiation factor IF-2 -
Bacillus subtilis >gp:gp|M34836|BACPSIF2A_1 B.subtilis prote
in synthesis initiation factor 2 (infB) gene, complete cds.
NID: g143358. >gp:gp|Z18631|BSORF1T7A_5 B.subtilis infB-nusA
operon. NID: g49314.

40

atgagtaaaaaaagaatttacgaatatgcgaaagaattaaatctaaagagtaaagagatt
atagatgagttaaaaagtatgaatgttgaaagtgtcaaatcatatgcaagctttagaagaa
gaacaaatcaaagcattagataaaaaatttaaagcctctcaagcgaaagacactaataaa
caaaatactcaaaataatcaccaaaaatctaataataaaacaaaattctaacgataaagaa
aaacaacaaagtaagaataatagtaaaccaacgaagaaaaaagaacaaaacaacaaagga
aaacagcaaaaataaaaaacaataaaaactaataagaatcaaaaaacaataaaaaataaaaag
aataataaaaaataataaacctcaaaatgaggtagcagaaacaaaagaaatgccctctaaa
atcacttatcaagaaggcataactgtcggtaggttagctgaaaagctaaatgtagaatca
gctggtattattaaaaaattgttcttactaggtattatggctaataatcaatcaatcattg
gatgaagaacattagaattaattgcagatgactatggcgttgaaatagagaaagaagta
gtcgttgatgaagaagatttatcaattttttgatgatgagactgatgattctgatgca
attgaacgtccagcagttgttacaatcatgggccacgtagaccatggtaaaacgacttta
ttagattctattcgtaaactaaagttacagaaggagaagctggcggaatcactcaacat
attggtgcttatcaaatgaaaattcaggtaaaaaattacgttcttagatactcctgga
catgctgcatttacgactatgcgtgcacgtggtgctcaagttactgatattacaatttta
gtcgtggccgctgatgatggtgtgatgcctcaacaattgaagctataaatcacgctaaa

55

gaagcagaagtacctacgattgttgagtaaacaaaattgataaaccaactgtaaccct
gatcgtgttatgcaagaactcactgagtgatgattaattccagaagactggggcggtgac
acaatctttgtaccactatctgcattgagtgagacggtattgatgatttattagaaatg
5 atcgggttagtagcggaggtacaagaacttaagctaatacctaataaacaagctgtaggt
actgtgattgaggtgaattagataaatcacgaggtccagctgcatctttacttgttcaa
aatggtacttttaacggttgagatgcaattgttgtaggtaatacttatggacgtatacgt
gcaatgggttaatgatttaggaaaaagaattaaatctgccggctccttaacacctgtagaa
10 attactggtatttaacgatgttccacttgaggtgatcggtttgttgatttgggtgatgaa
aaacaagcacgtcgaattggtgaagcacgtcatgaggcaagtgtcatacagcaacgtcaa
gaaagtaaaaatgtttcattagacaatttatttgagcaaatgaaacaaggtgaaatgaaa
gatttaaatgtcatcattaaaggtgatgtacaaggttcagttgaagcattggccgcatct
ctaataagaaatagatgttgaaaggtgtgaatgtacgaattattcatacagctgttggtgct
atcaatgaatcagatgttacattagcaaatgcatcaaatggtattattattggtttta
15 gtacgcccagatgcaggtgcgaaacgtgccggtgaagctgaaaatgtagatatgcgatta
cacagagttatctataatgttattgaagagatagaatcagctatgaaaggtttacttgac
ccagaatttgaagagcaagtcatggacaagctgaagtgctcaaacattttaaagtttct
aaagttggtacaattgctggttagttatgtgactgaaggtaaaatcactcgtaacgctggt
gtacgcgtaattagagatggtatcggtgttatttgaaggtgaacttgacacattaaaacgt
ttcaagatgatgctaaagaagtagctcaaggctatgaatgtggtattacaattgaaaa
20 tataatgatctcaaagaaggagacattattgaagcggttgaaatggtagaaattcaaaga
taa

Sequence 2058

MSKKRIYIYAKELNLKSKEIIDELKSMNVEVSNHMQALEEEQIKALDKKFKASQAKDTNK
25 QNTQNNHQKSNKQNSNDKEKQSKNSKPTKKKEQNNKGKQNNKNNKTNKNQKNNKNNK
NNKNNKPQNEVAETKEMPSKITIQEGITVGELEKLNVESAGIIKKLFLLGIMANINQSL
DEETLELIADDDYGVIEKEVVDEEDLSIYFDDTDSDAIERPAVVTIMGHVDHGKTTL
LDSIRNTKVTEGEAGGITQHIGAYQIENSCKKITFLDTPGHAAFTTMRARGAQVTDITIL
VVAADDGVMQPTIEAINHAKAEVPTIVAVNKIDKPTANPDRVMQELTEYGLIPEDWGGD
30 TIFVPLSALSGDGDIDDLLEMIGLVAEVQELKANPNKQAVGTVIEAELDKSRGPAASLLVQ
NGTLNVGDIAIVVGNTRYIRAMVNDLGKRIKSAGPSTPVEITGINDVPLAGDRFVVFGE
QARRIGEAREHASVIOQROESKNVSLDNLFEQMKQEMKDLNVIKGDVQGSVEALAAS
LMKIDVEGVNVRRIHTAVGAINESDVTLANASNGIIGFNVRPDAGAKRAEAENVMRL
HRVIYNVIEEIESAMKGLDPEFEEQVIGQAEVRQTFKVSQVGTIAGSYVTEGKITRAG
35 VRVIRDGIVLFEGLDTLKRFKDDAKEVAQGYECGITEKYNDLKEGDIIEAFEMVEIQR
*

Sequence 2059

Contig_702_pos_1521_1138,
40 is similar to (with p-value 2.0e-34)
>sp:sp|P32731|RBFA_BACSU RIBOSOME-BINDING FACTOR A (P15B PRO
TEIN). >pir:pir|G36905|G36905 protein P15B homolog - Bacillu
s subtilis >gp:gp|Z18631|BSORF1T7A_7 B.subtilis infB-nusA op
eron. NID: g49314. >gp:gp|Z99112|BSUB0009_135 Bacillus subti
45 lis complete genome (section 9 of 21): from 1598421 to 18072
00. NID: g2633902.
gtgagtaaaagtaaaaaataaagagaggtgagatgatgaataatataagagcagaacgtgta
ggagaacaaatgaaacaggaaatcatggacattgttaataataaagttaaagaccctaga
gttggttttttaacaattactgatgttgaaactaaccaatgaccttcacaagcaaaggta
50 tatttaacagtggttagggaatgataaagaagttgataatacgtttaaagctttgcataaa
gcaactgggtttataaaatctgaacttggttctcgaatgagcctaagaattatacctgag
ttaacattcgaatatgatgaatctatcgaataacggttaataagatagaacgcattgattcaa
gagttacacaaaaatgataaataa

Sequence 2060

VSKVKIKRGEMMNNIRAERVGEQMKQEIMDIVNNKVKDPRVGLTITDVELTNDLSQAKV
YLTVLGNDKEVDNTFKALHKATGFIKSELGSRMLRIIPELTFEYDESIEYGNKIERMIQ
ELIKQDK*

Sequence 2061

Contig_0702_pos_1000_83,

is similar to (with p-value 6.0e-46)

>gp:gp|X92946|LLLPK214_14 Lactococcus lactis sp. lactis plas
 5 mid pK214, complete sequence. NID: g2467210.
 atgtataatggcatactaccggtattttaagaaacgaggtttaacaagtcacgacgctcgtt
 tttaaattacgtaaaattttaaaaaatgaaaaaattggtcatacaggaaacattagatcct
 gaagttaatggtgtgtttaccaatttggttaggcgatgcgacaaaagtcagtgattatc
 atggaaatgggaaaaacttatcacgctatgataacgctaggaaagagtacgactactgaa
 10 gaccaaactggagatatttttagaaactagggtgttgataagaatgatattaatgaagat
 acgattgaccaagtgttgacgaatttgaggggcatattcaacaaattccgcctatgtat
 tcttctgttaaaatgaaatggaagaaaatttatgaatatgcgagaaaataatgaaactgtc
 gaacgccctaaacgacaagttttttattaaagatatacatagaatatctgaagttactttt
 caggagcagacatgtcattttgaagttgaagtaacatgtggtaaaggaaacttatattaga
 15 acttttagctacagatattggacttaaaacttggtttccagctcatatgtcacgtctaact
 agaattgcttctggcggttttcaattagaaagtagtttaacgattgatcaaattaaagaa
 ttacatgagcatgattcattacataatgaattgttccctatagaatatggcttaaaaggt
 ctgaaatcattccaagtgaagattcaaattttaaaaagaaaatctgtaacggtcaaaaa
 ttcataaaaaaagtggttaagtcaaaatgttaagaaccttttatatttgcgatagtagc
 20 actcaaaaagtttttagcaatatatatagttcatccagataaaccttatgaaataaaacct
 aaaaaagtttttaattaa

Sequence 2062

MYNGILPVFKKRGLTSHDVVFKLRKILKMKKIGHTGLDPEVNGVLPICLGDATKVS DYI
 25 MEMGKTYHAMITLGKSTTTEDQTDILETRAVDKNDINEDTIDQVLQQFEGHIQQIPPMY
 SSVKVNKRKLYEYARNNETVERPKRQVFIKDIHRISEVTFQEQTCHFEVEVTCGKGTIIR
 TLATDIGLKLGFPAHMSRLTRIASGGFQLESSLTIDQIKELHEHDSLHNELFPIEYGLKG
 LKSFQVKDSNFKKIKCNGQKFHKKVLSQNVKEPFI FVDSSTQKVLAIYIVHPDKPYEIKP
 30 KKVEN*

Sequence 2063

Contig_0703_pos_784_1236,

putative peptide of unknown function

atgtatgatgaattagagggttaataaaaagtcgttttaaaaactgtaattttaatgaaggt
 35 atttttaagaatatagaagcaatttgtaattgtaaaattacaacgtgcgggtttaataat
 tgtattttcgaagatgttcatttttcaaaaaaccaatttaagattcaacatttgatgaat
 acaccatttgatcaatccgtatttaatatgtactttattccaaaatgcaatgttcgatagc
 aatctcattcgtagcgtaaaatggactgatatacatttttaaaaacggttctttcaaaaat
 gtagaattgaaaggaacaacatttaagatgtaaaattcaaaaattgtgagttcaaaaat
 40 gtaattattactaattcaactatgtcgcgaaaagttaatgaatgaattacaaaaacaagat
 gttacttttagaaaatatagacacttctatttaa

Sequence 2064

MYDELEVNSRFRKNCNFNIEGIFKNIEAICNCKFTTCGFNNCIFEDVHFYKNQFKDSTFVN
 45 TPFQSVFNSTLFQNAMFDSNLIRSVKWDIIIFKNVSFKNVEIEGTTFKDVKFKNCEFKN
 VIITNSTMSQKLMNELQKQDVTLENIDTSI*

Sequence 2065

Contig_0703_pos_4474_5274,

putative peptide of unknown function

atgaaagatagggttgaaacagaagaatatgctagaaatcaattaatctctaaaaattca
 50 attttaagtgaagaaaatttatcattgaaaaccaaattgtaagtacaaaacaatgacgtc
 ggtcaaacacgcttttaaaaacgccaagcgtgaatttaagaaaaatattaaatagatttaa
 gaagaggggtcgtttacgatcatatacaattgttctacgagtaatttggtgtgtaaacat
 55 ccccttttcgaatatgcacgttcattcgattttattatcattactgatgttggttgata
 aatgtggatgttaaaaattggaacaaaaaacgttttatcattttgatgtgccagatcaa
 catcttgaagaaggacaaccacaatataataccgaaaaagttgtcgggtcattatattagc
 aatcgatatcatagtcagtttaaaaacaacacggttctggtgtctatacttttattgagatt
 ttacaggataatcgtgtaatatatgaattttatgaccacgatccatacgataaagccgca

aacaatgcaaaagcattaaaagataaaattgaaaatgattataattttaaaattcaaagt
 attggcgatcatatatttttagtgatggtagcggttaatatattgaaggatccgacgagagt
 gataaatacgtcgacaccggtatctacaccgatataccttgaaaaagtaattgaggaagct
 atcgatttatctaagcacccttactgataaacaatcgaagaatttctgaaaacttt
 5 aaacaacatatgaataattaa

Sequence 2066

MKDRVETEEYARNQLISKNSILSEENLSLKNQMLSTNNDVGQHAFKNAKRELKILNRFK
 EGRRLRSYITIVPTSNLAVKHPLFEYARSFDFIITDVGLINVDVKNWNQKTFYHFDVPDQ
 10 HLEEGQPQYNTEKVVGHYISNRYHSQFKTTRSGVYTFIEILQDNRVIEFYDHPYDKAA
 NNAKALKDKIENDYNFKIQSIGVIYFSDGSVNIIEGSDSKYVDTVSTPISLEKVIIEA
 IDLSKHPLTDKQIEEISENFKQHMMN*

Sequence 2067

15 Contig_0703_pos_3987_3481,
 putative peptide of unknown function
 atgcatctatttctgaatgatgattttatattaccatcgaaaatagatatttgtgttgag
 cgcaaaaagttattgaaattatcaggggtgattcctgaagaatttacgattcactattat
 cgaatcagttatataagagattaagagaatctttatctttatcagatggtgaacatggt
 20 aaagtctttaaaaacgatacagaagtcagacaatatttgtttatgatggtgtaaatgat
 gaatggctatttaggttagatcatcatatacgtttacaaaaaataatatattttcat
 tctttaagttggaacgtagattatattaagccagaaatagttcttatgtatgatttaag
 agtgaacaaaaatatcatcagtttagtaattataaagctggtattgattctcttagttat
 tatcaattctatattttaaaattggtagtaggtgaacagcgtattaaaaaagctatagta
 25 aatagttccactaaaaagatatcttaa

Sequence 2068

MHLFLNDDFILPSKIDICVERKKLLEIIRVIPEEFTIHYDNQLYERLRRESLSLSDVEHV
 KVFKNDTEVMTIFVYDVVNDEWLFRLDHHIRLPKNNIYFHSLSWNVDIKPEIVLMDL
 30 SEQKYHQFSNYKAVIDSLSYQFYILKLTVGGEQRIKKAIVNSSTKKIS*

Sequence 2069

Contig_0703_pos_2163_1420,
 putative peptide of unknown function
 35 atgtcgtcattatgccttttgcctggagctgaaacgataatgattttacgttcagggctc
 tcattcactataggacaatcgttttttacgctgtccctaggtacaaccggaatgattact
 tatgcaagctatgcacctaaaaatagacgataaaagcttcagcactttcaattgtcgta
 atgaatattttaatttctgtcttgctggattagctatatttctgcgcttaaaacattt
 40 ggttaccaccccaagaaggccctgcttattatttaaggtttaccactagtatttagc
 gaaatgacttttggtacattcttttactttatatttttactattattcttatttgcggca
 ttaacgtcttctatatcattattagagttaaatgtatctaattttactaaaaatgataat
 agtaaaagacaaaaagtggaatcataggtagtatacttgatttatcattagtatccca
 gcaacattatcttttagtagtctaagtcatttgcgttttggcgctggtacgatatttgat
 aatattgattttattgtatctaattcttatgccattaggggcactaggaacaacatta
 45 gtggttgccaattactagataaaaaattattaaaagaaagctttgggaagacaaattc
 aacctattttaccgtggtattatttaattaagttcatcatgcctattgttattatttta
 gtatttatagttcaattatttttaa

Sequence 2070

50 MSSLCLLPGAETIMILRSGSSFTIGQSFFTLSTLGTGMITYASYAPKNMTIKSSALSIVV
 MNILISVLAGLAIFPALKTFGYQPQEGPGLLFKVLPLVFSEMTFGTFFYFIFLLFLFAA
 LTSSISLLENVSNFTKNDNSKRQKVAIIGSILVFIISIPATLSFSSLSHLRFGAGTIFD
 NMDFIVSNILMPLGALGTTLVVGQLLDKKLLKESFGKDKFNLFLPWYYLIKFIMPVIL
 VFIVQLF*

Sequence 2071

Contig_0703_pos_0_399,
 is similar to (with p-value 9.0e-32)
 >gp:gp|U93874|BSU93874_1 Bacillus subtilis cysteine synthase

(yrhA), cystathionine gamma-lyase (yrhB), YrhC (yrhC), YrhD (yrhD), formate dehydrogenase chain A (yrhE), YrhF (yrhF), formate dehydrogenase (yrhG), YrhH (yrhH), regulatory protein (yrhI), cytochrome P450 102 (yrhJ), YrhK (yrhK), hypothetical protein YrhL (yrhL), putative anti-SigV factor (yrhM), RNA polymerase sigma factor SigV (sigV) and YrhO (yrhO) genes, complete cds, and YrhP (yrhP) gene, partial cds. NID: gl934604. >gp:gp|Z99117|BSUB0014_206 Bacillus subtilis complete genome (section 14 of 21): from 2599451 to 2812870. NID: g2634966.

atgattgcatacagatttgataggacaaactccattagttttattagaaagcttttagtgac
 gagaatgttaaaatatacgcacaaacttgagcaatttaacacctggtgtagcatcaaagac
 cgtctaggggaagtacttaattgaaaaagcaatagatgaaggacgacttaagaagggat
 acaatagtcgaagcgactgctggaatacaggcattggacttgctattgcttctaactcg
 cacaaagttaaagtatcatctttgctccagaaggatttgcagaagaaaaatttcaatt
 atgaaagcattgggtgcagatgttagacgtacccccaaagctgaggggaatgactggcgca
 cagcaagagcggttgccatacgcacacgatatgTACTA

Sequence 2072

MIAYDLIGQTPLVLLESFSDENVKIYAKLEQFNPGGSIKDRLGKYLIEKAIDEGRLKEGD
 TIVEATAGNTGIGLAIASNRHKVKCIIFAPEGFEEKISIMKALGADVRRTPKAEGMTGA
 QQEALAYATRYVL

Sequence 2073

Contig_0704_pos_804_1115,
 putative peptide of unknown function
 gtgttatacagaaaaatggttgctaaaaataaaacttaaacgcctatacaaacctattca
 caaactttatTTTTGTTGTTCATTTTAAATATTTTACTATTTTACTATAATCCTTGTT
 TTTTATTCCAAATACTTTACATCATCCTTGGGTAAGAAGGATTGATTCTTATCCTCATT
 AATAATAAATGTAATTATAAAGCCTTTCGCTGAAGTCAATAAGTCTGAATTCTAAAAAG
 CGAAACAGAAATCTTATCATATCTTCTTTGTTTCAATCCATTACCAGAACCCCACTTGC
 TTTGCTCTGTGA

Sequence 2074

VLYRKMVAKIKLRLYKPYSQTLFLLFILKYFYFTIILVFYSKYFTSSLGKEGLILILI
 NNKCNYSLSVNSISLNSKKRNRNLIIFFFVSIHYRTPTCFVC*

Sequence 2075

Contig_0704_pos_2016_0,
 is similar to (with p-value 3.0e-76)
 >sp:sp|P47994|SECA_STACA PREPROTEIN TRANSLOCASE SECA SUBUNIT
 . >pir:pir|S47149|S47149 secA protein - Staphylococcus carno
 sus >gp:gp|X79725|SCSECA_2 S.carnosus (TM300) secA gene. NID
 : g499333.

atgacctgtttatagatagaataactggctcgatgctacctggaacaaagcttcagtctggt
 ttacatcaagctatagaggctctggaaaatgttgaaatttctcaagatatgagtgtgatg
 gcaaccataacattccaaaacttatttaagcaatttgatgaattttcagggtatgactgga
 acaggtaaattaggggaaaaagaattctttgatttatattcaaaagtgttatagagatt
 ccgactcacagtcggattgaacgagatgatagacctgatagagtatttgctaattggtgac
 aaaaagaacgatgcaatttttaagacagtgattgggtatacatgaaactcaacaacctgtg
 ttactaattacacgtactgcagaagcggcagaatatttttcagctgagttatttaaacgt
 gatatacccaacaattttattaatcgctcaaaatgtagctaaagaggcacaatgatgct
 gagggcgacaaattatctgcagttactgttgctacaagtatggcaggcggtggaattgat
 ataaagttatcaaaagaggttcgatgatcggtggcttagcagtgattattaatgaacat
 atggataatagccgtgttgatcgctcaattaagaggacgctcaggtcgccaaggagatcct
 ggatattcacagattttttgtatcacttgatgatgatttagtaaaacgttgagtaactct
 aacttggcagaaaaataaaacctccaaacgatggatgcatctaaactagaaagtagtgca
 ctctttaaaaaacgtgtaaagtcaattgttaataaagcgcaacgtgtatctgaagagact
 gctatgaaaaatagagaaatggcaaatgaattcgaaaaaagtattagtgttcaacgagat

aaaatttatgctgaacgtaatcacatacttgaagcaagcgattttgatgattttaatttt
 gaacagcttgacagagatgtgtttacaaaagacgttaaaatcttgacttaagtagtgaa
 cgtgcacttgtgaattatatatacgaaaacttaagttttgtcttcgatgaagatgatca
 aatattaatatgcaaaatgatgaagaaatcatacaattcttaatacaacaatttactcaa
 5 caatttaacaatcgtttagaagttgctgctgattcatatttaaaacttcgtttcattcaa
 aatcaattttgaaagcgatagatagcgaatggattgaacaagtagataacttacaacaa
 cttaaagccagtgtaaacaaatcgacaaaatggacagcgtaatgtcatttttgaatatcat
 aaagtggtctcttgaaacgtatgaatatatgtctgaagatataaaaaggaagatggtaga
 aatttatgtttaagtattctagcctttgataaggacggagata

10

Sequence 2076

MTCLDRITGRMLPGTKLQSGHLHQAIEALENVEISQDMSVMATITFQNLFKQFDEFSGMTG
 TGKLGEKEFFDLYSKVVEIPIHSPIDERDDRPDRVFANGDKNDAILKTVIGIHETQQPV
 LLITRTAEAAEYFSAELFKRDI PNLLIAQNVAKEAQMIAEAGQLSAVTVATSMAGRGTD
 15 IKLSKEVHDIGGLAVIINEHMDNSRVDRQLRGRSGRQDGPYSQIFVSLDDDLVKRWSNS
 NLAENKNLQTMASKLESSALFKKRVKSIVNKAQRVSEETAMKNREMANEFKXISVQRD
 KIYAERNHILEASDFDDNFELARDVFTKDVKNLDSLSSERALVNYIYENLSFVFEDEVS
 NINMQNDEEIIQFLIQQFTQQFNNRLEVAADSYLKLRFIQKSILKAIDSEWIEQVDNLQQ
 LKASVNNRQNGQRNVIFEYHKVALETYEYMSEDIKRKMVRNLCLSILAFDKDGDGDX

20

Sequence 2077

Contig_0704_pos_1840_1244,

is similar to (with p-value 3.0e-87)

>sp:sp|P24277|RECR_BACSU RECOMBINATION PROTEIN RECR. >gp:gp|
 25 D26185|BAC180K_85 B. subtilis DNA, 180 kilobase region of re
 publication origin. NID: g467326. >gp:gp|X17014|BSRECM_3 Bacil
 lus subtilis dnaX and recR genes and two unidentified readi
 ng frames. NID: g453238. >gp:gp|Z99104|BSUB0001_21 Bacillus
 subtilis complete genome (section 1 of 21): from 1 to 213080
 . NID: g2632267.

30

atgcattatccagaacctatatcaaagcttatcgatagttttatgaaactgccaggcatt
 ggaccaaagacggctcaacgtctggcttttcatacttttagatatgaaagaagacgatgtt
 gtttaagtttgctaaagcactagttgatgtttaaagagaacttacctattgtagtgtttgt
 gggcatattacagaaaatgatccttgttatatatgtgaagataaacagcgagatcgttct
 35 gtcatatgtgtagttgaagatgacaaggatgtcatagcaatggaaaaatgcgtgaatat
 aaagggtttatatcacgtgcttcattggttcgatttcaccaatggatggtattgggcctgaa
 gacatcaatatacctgcattagttgaacgcctcaaaaacgatgaggtgaaagagcttata
 ttagctatgaatcctaactagaaggcgagctctactgcaatgtatatatctaggttggtt
 aaaccaattgggattaaagtcacaagactggcacaagggtttatctgtaggcggcgattta
 40 gaatatgctgatgaagtgactttatctaaagcaattgcaggtagaacggaatgtaa

Sequence 2078

MHYPEPISKLI DSFMKLP GIGPKTAQRLAFHTLDMKEDDVVKFAKALVDVKRELTYCSVC
 GHITENDFCYICEDKQDRSVICVVEDDKDVIAMEKMRKYKGLYHVLHGSISPMDGIGPE
 45 DINIPALVERLKNDEVKELILAMNPNLEGESTAMYISRLVKPIGKIVTRLAQGLSVGGDL
 EYADEVTLSKAIAGRTEM*

Sequence 2079

Contig_0706_pos_183_1418,

is similar to (with p-value 0.0e+00)

>sp:sp|P55189|YBAR_BACSU HYPOTHETICAL 46.4 KD PROTEIN IN RRN
 G-FEUC INTERGENIC REGION. >gp:gp|D84213|BACTHRTRNA_3 Bacillu
 s subtilis genome, trnI-feuABC region. NID: g1256147.

50

gtggcgcttttagtcacaccttttagttaagaacatgggtgtgaaatattcttttgctgct
 55 acgatattgatgggggttaatacaattacttttaggaatacttaaaagtcggtcgtttaatg
 aaatttatcccccgtccagttatgattggatttgtaaatgcattgggtattatgattttt
 atgtctcaaataagacatatttttaatatatctatagcaacttatatatacgttattata
 actttactaattgtatatgttattcctagattttataaaagctatacctgccccattaata
 gctataattgttttaacagcattgtatatgtatacaggatctgacgtaagaactgtaggt

gatttaggtaatattaagcaaaactttgcctcatttcttaatccctaacattccttttaatt
 ttagaaacacttcaaattattttcccatattcattatcaatggctattgtgggtctagta
 gaaagcctacttacagctaaaattgtagatgatgcaactgatacacacagtagtaaaaaac
 aaagaatcaagagggtcaagggtatgctaataattattactgggttctttggcggtatggga
 5 ggttgtgccatgattggacaatcagtaataatgtaagtcaggtgcaaatagtagattg
 tctacatttacagcagggtattgtacttatgtttatgattattgttctcggtggtgtagtc
 gttcaaatcccgatgccaatcttgcagggataatgggttatgggttctatcgggactggt
 gatttgaattcatttaagtataaaaaaagcacctaaaacagatgcattcggttatgggt
 ttaacggttattattgtatttaataaacacataacttagctatcggtgtagttgtaggtggt
 10 gttttcagtgctatattctttgcgactaaaatttctaaagttgaagttatttataaagag
 ttaggtaagcagcatcggttttcttttaagggtcagatattctttgtagtaattgattct
 atgatggagaaaattgattttaatatgaagatagtgtagtgaattttgatcac
 gctcacctatgggatgattcagcagtaaatgccattgatacgattgtaggaaatttgaa
 gagaaaaacaatattgtgtatgttgaaaaattaaacgcagatagccgaaaaattatttca
 15 gaactaagttattttaaatgaaactcattttaattaa

Sequence 2080

VALLVTPLVKEHGVEYLFAATILMGLIQLLLGILKVGRMLKFIIPRPMIGFVNALGIMIF
 MSQIEHIFNISIATYIYVITLLIVYVIPRFYKAIPAPLIAIIVLTALYMYTGSQVTVG
 20 DLGNIKQTLPHFLIPNIPFNLETLQIIFPYSLMAIVGLVESLLTAKIVDDATDTHSSKN
 KESRGQGIANIITGFFGGMGCMIGQSVINVKSGANSRLSTFTAGIVLMFMIIVLGGVV
 VQIPMPILAGIMVMVSIQTVDWNSFKYIKKAPKTDAFVMVLTVIIVLITHNLAIQVVG
 VFSAIFFATKISKVEVIYKELGKQHRFSFKGQIFFVSIIDSMMEKIDFNIEDSVIPLNFDH
 AHLWDDSAVNAIDTIVRKFEKNNIVYVEKLNADSRKIISELSYLNETHLN*

Sequence 2081

Contig_0706_pos_1422_1844,

putative peptide of unknown function

gtgttttttatgtataattcaattttgcttgcagcagatgggtcaaaaaatagtatacgt
 30 gcagcccaagagttactaaattttataggtgattatacaattgtcacaatacttacggtt
 gtagatattgaagaatcaaaaacagatgttttcatgatcaccaggaactaatctaact
 caaaaaagggaagttaattacaatctattaaagattttatttacagaacacaatgtgaat
 taaaaaattaaaattgttcacggtactccaacagataaagtgggtgaagtttcaaattagt
 ggtgaatatcaagctatcatttttaggtacacgtggcttaaacagttttcaaaaaatgggt
 35 cttggttagtgtaagtcataaagttgctaaccgttctcaaatacctgtaattattgttaaa
 taa

Sequence 2082

VFFMYNSILLAADGSKNSIRAAQELLNFIDYITIVTILTVVDIEESKTDVLHDHQGTNLT
 40 QKRESKLOSKDLFTEHNVNYKIKIVHGTPTDKVVEVSNSGEYQAIILGTRGLNSFQKMV
 LGSVSHKVAKRSQIPVIVK*

Sequence 2083

Contig_0706_pos_5297_6436,

putative peptide of unknown function

atgaaaaaatatggattttaactataggtatgtttgccttaggtatggatgcttatata
 gtagcaggatttaataccttcaataagtaaaagttttaataaaagttagctctgctattggg
 caaggagtaacagtttttacattgttttctctatctctgccccattttttcaacaata
 ttagctaaatccccagttaaaaaataactaataatagcattcagtatatttacttttagcc
 50 aatattataaccgcaatatctatgaactacatgctatatatcgatcaagagcaatcgct
 ggttttaggagctggcgatttctaccaattgcaataagtgaagcaatcatttagtctcc
 gaaaagcataaaggaaaagcaatcgcttttacagtaggcggaatgagtgtaggaactggt
 ataggagttcctctcgactagaaaattgccaacatttctaattggcgatttgcaatgttg
 gttattattgtcattagttttattgcattaataagcatatctatattgatgcctaaattt
 55 aatattagagctcctccaaatttaaaagatcgttttcaattatttttaacaagcatgta
 ctaagagttatttcgggttacattatgcgctgccattgctagtttaggtttgtatacttat
 ttagccgatattattaaaacaaatacagatacaaaaaatttaactcattaccttacagcg
 tggggaataggcggattaataggaagttttggtataggatttattatagatagatttaaa
 aatacaagatttgttatgctaattattttaattttactagcattaagttttgggttaatt

cctattttctattaacttgcctatatattaggtttaattccctttattttatggggagctatg
ggatgggctacacaagctcctcaacaacatatattattgaaaaacatcctgaatatgga
ggctctgctgctgcttttaaatagttctattaattatttaggcagtgctatgggatcagca
atcggaggaaattattttatttaagtctaatagtacaaatgtactaatatatagtgcttta
5 ggaattactattattggtattttattacaattactaaatttatccctagaaaaaaattaa

Sequence 2084

10 MKKIWILTIGMFALGMDAYIVAGLIPSISKSFNKSSSAIGQGVTVFTLFFSISAPIFSTI
LAKSPVKKILIIAFSIFTLANIITAI SMNYMLYIVSRAIAGLGAGVFSPIAISASNLVS
EKHKGKAIAFTVGGMSVGTIVGVLGLEIANISNWRFAMLVIIIVISFIALISISILMPKE
KIEAPPNLKDRFQLFLNKHVLRVISVTLCAAIASLGLYTYLADIKTNTDTKNLTHYLTA
WGIGGLIGSFGIGFIIDRFKNTRFVMLIILILLALSFLGIPISINLPILGLIPFILWGAM
15 GWATQAPQQHILLKKHPEYGGSAVALNSSINYLGSAMGSAIGGIILFNANSTNVLIYSAL
GITIIGILQLLNLSEKN*

Sequence 2085

Contig_0706_pos_5156_4323,
putative peptide of unknown function
20 atgggggattttatgaattttaattatatgaaggacttcattaaagttgtcgagtacaat
agcttaataaaagcttctagagaattgaatataagtaacaccagccttaagtaaaagaatt
agaagtattgaagattattttgattgtcagttattttatagaacttctaagggtattttt
ttaactcaaaaaggaaatgctgtttatcattcgtttttaaaaaataaacgagcaatttgaa
gaattaaaaagtaaaataagtgagtctaaagataaaagaattaaattaggagtcataacct
25 agcttttctttatataagctacatgaaaagaattatattaatgaaagtgctgttttagtt
attgaaaatagctacttcaatttttacttgaggaaaatttataaaagataatatagatgttggt
atttggtatttacacatttaaaaaataactcgtctctatacacagaattttacacagag
gactttattgtggttatggtgatgaaaataagtttaaaaacactaaccaagtaagtata
aaatctttaaaaaatgagaatatattatatacaaacaccgcatgtgatacatatgatctt
30 cttaaaaattattccttgaaagataagttacataattgattatgttgattattatgaaact
atattagctaattgtcaaagcaataaaggtattactttgactcctcaatctttgactaca
agacttgaaagtatgaacttatatcaaaaaaaattacaaagttataaaagagttgttaggc
gtagtatcacgtgataaagataaagatgaataaaattattgatataattcagtaa

35 Sequence 2086

MGDFMNFNFMKDFIKVVEYNLSLNKASRELNISTPALS KRIRSI EDYFDCQLFYRTSKGIF
LTQKGNVYHSFLKINEQFEELKSKISESKDKRIKLGVIPSPFSLYKLHEKNYINESVVLV
IENSTSILLEEYIKDNIDVIGDITHLKNSLYTQEITYEDFIVAYGDENKFKNTNQVSI
KSLKNENIYIQTPPCDTYDLKKNYSLKDKLHIDYVDYYETILANVKANKGITLTPQSLTT
40 RLESMNLYQKQLQSYKR VVGVS RDKDKMNKIIDIQ*

Sequence 2087

Contig_0706_pos_3655_3065,
putative peptide of unknown function
45 gtgggtgaattcaaattccctagtgatcgataaaagtaaacacgtaaaaggtattgaaaaa
gatcaggaagagattttaagacagtttgaaaaagaaaatgataaaagtatagaatttgat
gctgatgagatgagcgctaaaatcggagaattatttgagtagaatataaagatggttg
cctattgataattctggtggtggtggtgctccacctaaatcagaaagttataaccctca
gtttttaacgctcatctattttctatagtggttagtgagattctttaaaagaagagtat
50 aaatttttcggtttcaaagagaccattatctctttatatactgcctatttatctaattat
aaaaagataaagtggttactaaaggtggtgctccttatcatgtttcaaatagagaagag
agcttagagaagaaagcatttgatattataactgggtattatttagaagctatgaatact
cgtaagaatggaactttaatactataaaaatcttatcaaaagttaataagagttgtgttg
gtttcttatcaagttatagaacgagaacttatggaggaaaatgaagaataa
55

Sequence 2088

VGEFKFPSPGSIKVTTRKGIEKDQEEYLRQFEKENDKSI EFDAD EMSAKIGELFGVEYKDGL
PIDNSGGGGAPPKSESYNPSVFNAHLFSIVASEDSLKEEYKFFGFKETIISLYTAYLSNY
KKDKWVTKGAPYHVS NREESLEKKA FDIITGYYLEAMNTRKEWNFN TIKSYQKLIRVVL

VSYQVIERELMEENEE*

Sequence 2089

Contig_0706_pos_2458_2060,

5 putative peptide of unknown function
 atgcaaatgttacttaaagataaagttgacagtgaaggtaaattctattatctatgaagga
 caaagaggatctgggggaggagaattagatactcatgatcccatgtatttttcgaattta
 attactagtaaaatatttttaattaatcaagtagaaaaacataataagttttttatagct
 gaaaagctttttgaaggttatgaaagtaaatgtatggaaaaatatggaacaaattatcac
 10 ggagcaataatagagacaagagatgaaaataggatattagccgatgaaacaatacgtaaa
 ttttacttttttaacgaaagacatggaaaaagagaatgaatattttaagcttagatgattta
 aataaatttagactttttaatatttgcatttactcattga

Sequence 2090

15 MQMLLKDKVDSEGKSIIEGQRGSGGGELDTHDPMYFSNLITSKIFLINQVEKHNFFFIA
 EKLFEQYESKCKMEKYGTNYHGAIETRDENRILADETIRKFYFLTKDMEKENEYLSLDDL
 NKFRLLIFAFTH*

Sequence 2091

20 Contig_0707_pos_498_965,

putative peptide of unknown function
 atgcaattagatgtttttcctgaattaaacgtagatcaactatctcaaaaagtgagaaaa
 atacttaatgcagaaccagataaatatattaaaaatagcttgcgaggattaatagaagag
 cggattttattgtttattctcgaacaatctggtattaatgatgaaatgactgcacatcat
 25 ttatcaaatcaacagttttcaaaacttttattaatcttctaaagacttttacctttacagta
 gatggcacacttccattagacaaaagcttttgtagcaggcgcggtatttcttttaaagaa
 atagaaccaaaaactatgatgtctaaattagtgccaggattatttttatgtggcggaagtc
 ctagatattcatggatacactggaggttataatataacaagcgcattagttacgggtcat
 gttgctggaatgtttgctggtgaattttaaataatagatcaaaaacaaataa

30

Sequence 2092

MQLDVFPPELNVQQLSQKVRKILNAEPDKYIKNSLRGLIEERYLLFILEQSGINDEMTAHH
 LSNQQFQTFINLLKTFFTVDGTLPLDKAFVTGGGISLKEIEPKTMMSKLVPLGLFCGEV
 LDIHGYTGGYNITSALVTGHVAGMFAGEFKIDQNK*

35

Sequence 2093

Contig_0707_pos_1717_1337,

putative peptide of unknown function
 atgataaatgcaaatttaagtgttagtactagagtgaacaaaaatgctcgtagatcgata
 40 aatgaaattgttagtaacgctttaagtcaacttaataaagtaaccacaaataaagaagtt
 gatgaaatagtttaacgaaacgattgaaaaacttaagtcaatacaataaagagaagataaa
 atattgagtagtcaacgttcatcaacatctatgacggaaaaatctaataatgttatagt
 tccgagaataatacaattaaatctctaccagaggcaggaaatgctgataaatcactacca
 ttagcaggagttacttttaatatctggttttagcaatcatgtcctcacgtaaaaagaaaaaa
 45 gataaaaaagtaaatgactaa

Sequence 2094

MINANLSASTRVKQARTLINEIVSNALSQLNKVTNKEVDEIVNETIEKLKSIQIREDK
 ILSSQRSSTSMTEKSNQCYSENNTIKSLPEAGNADKSLPLAGVTLISGLAIMSSRKKKK
 50 DKKVND*

Sequence 2095

Contig_0707_pos_0_497,

putative peptide of unknown function
 55 atggatatttttctttttttgatttttctgctctttataaaacaaactgactactggagat
 gattcaatgaagacctataagccgtaccgacatcaattaaggcgttcgctatttgctca
 acgattttccagattttatggatgattattgggttaataagcttttatgctatttat
 atcgggtcgaacattgcaccattcatcagcatacctatcaaactcaaacccaatttcaa
 cgtatcgacaaacattttcatacgtttgttacgcagcaacaaaaacaatggcgtcatgtt

gatttatcacatccaactgatatcacaagatgaaacgccaactattaaacaagtcct
 caacaacctgcgatattgtattacgattttaaaggttcttcacaatctttcacaacaat
 tatgaacaattagacacaacaagatgtatttaatatcaaaatatcgaattgatttttaa
 gacgatacttatatcct

5

Sequence 2096

MDIFLFLIFLLFINKLTTGDDSMKTYKPYRHLRRLFASTIFPVFMVMIIGLISFYAIY
 IWVEHCTIHQHTYQTQTELQRIDKHFTFVTQQQKQWRHVDLSHPTDITKMKRQLLKQVH
 QQPAILYYDLKSSQSFTNNYQLDITKMYLISKYRIDFKDDTYIL

10

Sequence 2097

Contig_0708_pos_1581_3635,

is similar to (with p-value 0.0e+00)

>gp:gp|AF090142|AF090142_1 Staphylococcus epidermidis lipase
 precursor (gehD) gene, complete cds. NID: g3789931.

15

gtgatttctttgacaaataataatagaataaaaaagatttagcattagaaatacgcagtg
 ggagtagtatcaattattacaggtgtaacaattttcatcgaggggcaacaagctcaagca
 gctgaaacttcagtgcaacatgcggatgcacaccagaagactcgcaaaacaacacaacaa
 ttaaaaaatgataaggtagaagaacggttaaaaagcttcaaaacaaggtactgcttatagt
 caacaagtcaaaacaatttaattcaatctaaaacaattcaaaacaaccaacattctgtagct
 20 gaaagtgaacaactgaaaagtgtgagacagctaatcagccaaaaaagaagaaggtagt
 tcagtaaaacaagacgctccaaccatctaaaaatgtaaatcaacaagacgcagctactcaa
 tcaaatgagagaaaaaatattgacataaaaggtgaaggtcaaacttcaaagacaagcaat
 caacatattcagagttctaacagtcataatcaatcaacagaaacaaaagacagcgactca
 25 gaagaaatcgatcaaccattagtgaattacaaaagccgtctaatgattctacatatcaa
 acacaatcaaaagctaaacaagatagttctaaacagctccctcaagaaaaacaacaaaa
 cgtcaaatccaaacaactgaaaatgaacagacaactaaagttgattctaaaaaagctaat
 gacactcaaaatggtgaacaacatactcaagagcctaaaaatgatatacaacatcacaa
 aaaaatcatcatcaagtagctacaaaagaacaatctaatagaagtacaacaagggaagcg
 30 caaaagcgaatcagcaaatgctaatcaaaatcatcagttctacacatcaagcacagttcaaa
 aaccaatatccagtagtatttgtccatgggttcttaggctttgcaggtgataatcaattt
 agtttagctccaaaatattgggggtgtacaaaatacaatattgacagaaaatttaactaat
 gaggatatacatgtacatgaggcaaatattgggtgcttttagtagtaactatgatcgcgca
 gtagaattgtattactatgtcaaaggaggacgtgttgattacggtgcagcgcatgcagct
 35 aaatatggtcatcatcgctatggtcgaacatacaagggcacatgcgtgattgggaacct
 ggcaaaaaaattcattttataggtcacagtatgggtggtcaaaccattcgtcaaatggaa
 gaattcttaagaaatggtaaccaagaagaatagaatatcaacgtcaacatgggggtact
 atatccgatttatttacaggtggttaagataatatgggtgcttcaattactacacttggc
 acaccacataatggtacacctgctgcagataaaaattggcacacgtaaacttgtaaaagaa
 40 acgattaatcgatttgtagattagtggtggtgtaagatgtagatatagatttaggtttt
 tctcaatggggattaaaacaacaaccaaatgaaagctacattgattacgcggaacgtgta
 tccaaaagtaagatttggaaactgaagatcaagctgttaatgatctgacaacgcaaggt
 gctgaaaaaattaatcaacaacaagctctaaatcctaatttgtctacactacttataca
 gggtcagcgactcacacaggacctcttggttaatgaattacctaattctagtgaattctt
 45 ttggtgaacttaaccagccgtattattggtaagatgcaaaacaagaaattagaccgaat
 gatggtgtagttccagttatatcatcacacacaccttctaatacagcctttaaaaaagtt
 gatgatcatacaccagctactgataaaggtgttggcaagtgagaccggttcaacatgat
 tgggaccatttagatttagtaggtatggatgcatttgatttaacacatacaggtagagaa
 ttgggtcaattctatctaggtattatggataatatcatgagaatcgaagaagcagcagggt
 50 attacaaataataa

Sequence 2098

VISLTNNNRIRKFSIRKYAVGVVSIITGVTFIFGGQQAQAETSQVQHADAHPEDSQTQQ
 LKNDKVEETLKASKQGTAYSQQVQVINQSKTNQNNQHSVAESEQLKSDATANQPKKEEGS
 55 SVKQDVQPSKNNVQQAATQSNERKNIDIKGEGQTSKTSNQHIQSSNSHNQSTETKDS
 EEIDQPLVKLQKPSNDSTYQTSKAKQDSSKQLPQEKTTKRQIQTTENEQTTKVDSKKN
 DTQNVQHTQEPKNDTSTSQKNHHQVATKEQSNRSTTRETQKQSANANQNHQSTHQAQFK
 NQYPVVFVHGLGFAGDNQFSLAPKYWGGTKYNIDRNLNTEGYNVHEANIGAFSSNYDRA
 VELYYYVKGGRVDYGAHAHAKYGHHRYGRTYKIMRDWEPGKKIHFIHGHSMGGQTIRQME

5 EFIRNGNQEEIEYQRQHGGTISDLFTGGKDNMVASITTLGTPHNGTPAADKIGTRKJ,VKE
 TINRFGRLSGGKQVDIDLGFSSQWGLKQQPNESYIDYAERVSLSKIWNTEQAVNDETTQG
 AEKINQQTSLNPNIVYTTYTGSAHTGTPLGNELPNSSEILLNLTSRIIGKDANKEIRPN
 DGVVPVISSQHPNSQAFKKVDDHTPATDKGVWQVRPVQHDWDHLDLVGMDAFDLTHTGRE
 LGQFYLIGIMDNIMRIEADGITNK*

Sequence 2099

Contig_0710_pos_168_1115,
 is similar to (with p-value 7.0e-45)

10 >gp:gp|Z99108|BSUB0005_69 Bacillus subtilis complete genome
 (section 5 of 21): from 802821 to 1011250. NID: g2633055. >g
 p:gp|D78509|D78509_11 Bacillus subtilis YfjG-YfjR genes, com
 plete cds. NID: g2780390.
 atgacggtgaccgtagatttcaatctttatcgcaacctcttacattagtttcaaagtgtg
 15 aaagagattcctaagatgcaacgattatatggtatgattttgaaaatgccactgatgaa
 gaaaatgagtagtttaaaaaatcattttgatttcaattacttagaaatagatgatgctatc
 aatggtgacccacgagttaaatatatagaatatgacgcgtatcaatatatgatatttcat
 agtattattaatgatgattactcaccaatctcactaagtgtatttttagaaggtaagtgt
 ttagtgacataccatcacaaacattttccatcattaaagcgtgtggctcaatacaatgca
 20 gaaaatcatgatgattgaattagattgtgcagacatcgctcattcatattctggattgtatg
 gtggataaataattttaactttgtttatggtattgaagataaagtgtataattttgaagct
 aagcatgtcgatgaccgctatagtaagagcgttatggaaaatgtctttcaattacgttcg
 gatttaattaaaatcaaacgcgtattttccgatgcaagaagttgtagatacaatgaaa
 caagaaggaaatataattaaagatgccaacatagaatgtatattcaacatattgatgat
 25 catcttattaaacaaagaagtgttattcggacttctcaagaaatgacgaatgagattcgt
 gaaaattatgaatcattcacctcatttaggatgaatagtataatgcagatacttacgctt
 gtatctgttatattctcaccactcacttttattgctggtgtatatggaatgaactttgaa
 tttatgcctgagttgaaatggcattatgcttatttcgtgtgcttaactttaatgctaatt
 ataacaataatattaatcatattctttaaaaagaaaaaatggttttaa

Sequence 2100

MTVTVRFSQSLQPLTVSNVKEIPKDATIIWYDFENATDEENEYLKNHFDNYLEIDDAI
 NGDPRVKYIEYDAYQYMI FHSI INDDYSPI SLSV FLEGNVLV TYHHKHFP SLKRVAQYNA
 ENHDS ELD CADIVI HILDCMVDKYFNFVY GIEDKVYNFEAKHVDDRYSKSVMENVFQLRS
 35 DLIKIKRVLFPMQEVD TMKQEGNI IKDAKHRMYIQHIDH LIKQRSVIRTSQEMTNEIR
 ENYESFTSFRMNSIQILTLVS VIF SPLTFIAGVYGMNFEFMP ELKWHYAYFVCLTLMLI
 ITIILIIFFKKKKWF*

Sequence 2101

40 Contig_0710_pos_1206_1763,
 is similar to (with p-value 2.0e-29)

>gp:gp|AF016485|AF016485_60 Halobacterium sp. NRC-1 plasmid
 pNRC100, complete plasmid sequence. NID: g2822278. >gp:gp|AF
 016485|AF016485_149 Halobacterium sp. NRC-1 plasmid pNRC100,
 45 complete plasmid sequence. NID: g2822278.
 atgtcacaaaaagatgctggtttcagattttgataaagtgagatttgttcattcattcc
 atccccagtagttgatgttagtcaagtcgatatgacaagtcatactacgaaattcgatttg
 gcatatccaatctatataaatgcaatgactggtggaagtgtattggacaaaacaaattaat
 gaaaaatttagcaattgttgctagagaaactggaattgcaatggcgggtgggatcaacacat
 50 gcagctttgcgcaatcctaataatgattgaaacatttagcattgtgcgtaaaacaaatccc
 aaaggaacaattttcagcaatgtgggtgccgatgtaccagtggataaagctctacaagcg
 gttgcaattattagatgctcaagcgtacaaattcatgtgaactcacctcaagaattagtc
 atgcctgaagggaaccgtgaatttgcttcattggtatgtcaaatattgaatctattgttaaa
 cgcggtgatgttccagttattattaaagaagttggtttcggaatgagtaaaagaacatta
 55 caagcggttatatgattaa

Sequence 2102

MSQKDALVSDFDKVRVHHSIPSIDVSQVDMTSHTTKFDLAYPIYINAMTGGSDWTKQIN
 EKLAIVARETGIAVAVGSTHAALRNPNMIETFSIVRKTNPKGTFISNVGADVPVDKALQA

VELLDAQALQIHVNSPQELVMPEGNREFASWMSNIESIVKRVDVPVIIKEVGFMSKETL
QALYD*

Sequence 2103

5 Contig_0710_pos_1830_3305,
is similar to (with p-value 4.0e-90)
>gp:gp|AF006665|AF006665_31 Bacillus subtilis 168 region at
182 min containing the cge gene cluster. NID: g2529445. >gp:
10 gp|AF015775|AF015775_7 Bacillus subtilis YodA (yodA), YodB (yodB),
YodC (yodC), YodD (yodD), ABC-transporter (yodE), permease (yodF),
proteinase (ctpA), YodH (yodH), YodI (yodI), carboxypeptidase (yodJ),
purine nucleoside phosphorylase (deoD), YodL (yodL), YodM (yodM),
YodN (yodN), YodO (yodO), YodP (yodP), acetylornithine deacetylase (argE),
butyrate-acetoacetate CoA transferase (yodR), butyrate acetoacetate-CoA
15 transferase (yodS), YodT (yodT), CgeE (cgeE), CgeD (cgeD), CgeC (cgeC),
CgeA (cgeA), CgeB (cgeB), YzxA (yzxA), UDP-glucose epimerase (yodU),
YodV (yodV), and YodW (yodW) genes, complete cds; and YodZ (yodZ) gene,
partial cds. NID: g2415383. >gp:
20 :gp|Z99114|BSUB0011_121 Bacillus subtilis complete genome (section 11
of 21): from 2000171 to 2207900. NID: g2634230.
atgaatgatcatcaaaaaaatcatgcaacatctcaagatgataacacaatgtcaacacca
tctaagaatagcaagcatataaaaaattaaattatggcatttcatactcggtattttgggt
attattcttttaacatccatcattactgtagtatcaacaattttaattagccatcaaaaa
25 agtgggttttaataaagaacaacgtgcaaatttaaaaaaattgaatatgtctatcaaacac
cttaataaagattattacaaaaagcaagttctgataaattaactcaatctgccatagat
ggatgggttaaagaacttaaagatccatattcagaatatatgactgctgaagaacaaaa
caatttaataagagtgatcagggtgatttcggttgccataggtgctgaaatgcaaaagaaa
aatgaacagataagtggttactagcccaatgaaggattcaccagcagaaaaagctgggtatt
30 caacctaaagatatcgtcacacaagtgaatcatcattcggtagtcggtaaacacttgat
caagttgttaaaatgggtccgcggaaaaaaggaacatatgttactttaactataaaacgt
ggttcgcgaagaaaaggatattaagattaaacgcgataaccattcacgttaagagtgtagag
tatggaagaaaggcaatgtaggcgtactaacaatcaataaattccaagcaataattct
ggatgaactcaaatctgcaatcatcaaaagcgcataagcaaggcatccgtcatatcatttta
35 gatttgagaaataatccgggggggttattagatgaggcagtcagatggctaacatcttt
attgataagggaaatactgtcggttcaattagaaaaaggtaaggataaggaagaattaaaa
acttctaataagcactaaaacaagcaaaagatatgaaagtatccatcttagttaatgag
ggatcagctagtgcttcagaagtgttacaggtgctatgaaagactatcataaagctaaa
gtttacgggttctaaacatttggtaaaggtatcggtcagaccactcgatgaatttagtgat
40 ggttcattaattaaatatacagagatgaaatggctaacgcctgatggccattatattcat
ggtaaaggaattagaccagatgtagtatctcaacacccaaaataccaatcactcaatgtc
attccagataacaaaacttatcatcaagggtgaaaaagataaaaatgttaaaacgatgaaa
ataggtctaaaagcttttaggttatccaattgataacgaaacaaacatatattgacgaacaa
ttagaatctgctattaaaacatttcaacaagacaataatttaaaagttaatggcaatttt
45 gataaaaaaacaaatgataaatttactgaaaaactagttgaaaaagcgaataaaaaagat
actgttttaacgatttactaaacaaactaaaataa

Sequence 2104

50 MNDHQKNHATSQDDNTMSTPSKNSKHIIKLWHFILVILGIILLTSIITVTVSTILISHQK
SGLNKEQRANLKKIEYVYQTLNKDYKQSSDKLTQSAIDGMVKELKDPYSEYMTAEETK
QFNEGVSQDFVIGIAEMQKKNEQISVTSMPMKDSPA EKAGIQPKDIVTQVNHHSVVGKPLD
QVVKMVRGKKGYTYTLTIKRGSEKEDIKIKRDTIHVKSVEYEKKGNVGVLTINKFCUNTS
GELKSATLKAHKQGIRHIILDLRNNPGGLLDEAVKMANIFIDKGNVTVQLEKGDKEELK
TSNQALKQAKDMKVSILVNEGSASASEVFTGAMKDYHKAKVYGSKTFGKGIVQTTREFSD
55 GSLIKYTEMKWLTDPGHYIHGKGI RPDVSI STPKYQSLNVI PDNKTYHQGEKDKNVKTMK
IGLKALGYPIDNETNIFDEQLESAIKTFQDDNNLVNGNFDKKTNDKFTKLEKANKKO
TVLNDLLNKLK*

Sequence 2105

Contig_0710_pos_4812_3682,
is similar to (with p-value 7.0e-51)
>sp:sp|P53434|YRP2_LISMO HYPOTHETICAL 41.4 KD PROTEIN IN RPO
D 2' REGION (ORFA2). >gp:gp|U17284|LMU17284_3 *Listeria monocytogenes* major sigma factor (rpoD) gene, partial cds, and downstream orfA1 and orfA2 genes, complete cds. NID: g687597.

5 atgggttggaatgtaggattattaattggaatgataagttagatattacagggtattctg
acaacactcgactgcaccgatgatgttgtaaccaagcaattgaacttaataccaatacc
atcattgctcatcatccacttattttcaaaggagtaaaacgatatcggtgaagatggatat
10 ggtagtataattcgtaaaacttatccaaaataatatcaatcttatagcattacacactaat
cttgatgtaaactcctaaagggtgtcaatcgaatgtagcggatcaaatagggttagagaac
atatcaatgattaatacaaatagctcatattattacaaagttcaaactttatacctaaa
aattatattgaagatttcaaagacagtttaaacgaacttgatttagctaaagaaggtaatt
tacgaatattgtttctttgaaagtgaaggtaaagggaatttaaaccagtaggtgatgca
15 agtcccttatataggaagttagatagtagatcgaatatgttgatgaaataaaacttgagttt
atgataaaagacaatgaattagaaataactaaacgtgctatttttagataatcaccatac
gaaacaccagtttttgattttattaaaatgaacaaagaaagtgaagtatggattagggatt
attggacaattaaacaaaactatgacttttagatgaattttctgaatatgccaaaaaacag
ctcaatataccgagcgtagcatatacaggtcaacatgatagtcgaattaaagaagtagct
20 atcatagggtggttcaggtagataggattttagtataaagctagccaacttgagcagatgtt
tttggtactggtgatattaaacaccatgatgcttttagatgctaaaatccaaatgtaaat
ttattagacatcaatcattatagtgagtagttatgaaagaaggattaaaagaattatta
gaaaaatggttattttaaataatgaaaatcaatttccaatatatgcttctgaaatcaacaca
agtgttttattttcgatgtctttcttcaaaatgctctttcgcgggatattcatcatccat
25 tggttggttggaattcttccataatctggttagcctctttgactaattga

Sequence 2106
MVGNVGLLIGNDKLDITGILTTLDCTDDVNVQAIENLNTNTIIAHHPLIFKGVKRIVEDGY
GSIIRKLIQNNINLIALHTNLDVNPKGVNRLADQIGLENISMINTNSSYYKVQTFIPK
30 NYIEDFKDSLNLGLAKEGNEYECFFESEGGQGFQKPVGDASPIYIGKLDSEYVDEIKLEF
MIKDNELEITKRAILDNHPYETPVFDFIKMKESEYGLGIIGQLNQTMTLDEFSEYAKKQ
LNIPSVRYTGQHDSPIKKVAIIIGSGIGFEYKASQLGADVFTGDIKHHDALDAKIQNVN
LLDINHSEYVMKEGLKELLEKWLFKYENQFPIYASEINTSVLFMSSSKCSFAGYSSSH
WLLNSSIICLASLTN*

Sequence 2107
Contig_0711_pos_2494_2901,
putative peptide of unknown function

40 atgaaattatctattttcatgatgggtcaattttttgtaggtgctggtgaataccaagaa
ggcttcattcataaataatctaaaagttacatttggaatgaacctagcgatgaaacagt
ttacgattcataacttttaacttattcctttattaaatcaaacacacggttaagaagaaa
cctattcaaaagcataaaaagattaatccaaaacggtttacaacgtaaaatcgctaaagaa
caaaaagagaccaatttaactacatttgctcaacaagcgattaaagaagaacaagaattg
aataagctaaagagtaaaaaacttcagcgattagaaaaagaacgacacagacaatacaaa
45 agaattgttaaaaagaaaaaagcacatgaaaagcacaaggtcactaa

Sequence 2108
MKLSIFHDGQFFVGVVEYQEGFIHKYLKVTFGNEPSDETFLRFITFKLIPLLNQTHGKKK
PIQKHKKINPKRLQRKIAKEQKETNLTTFAQQAIKEEQELNKLKSKKLQRLEKERHRQYK
50 RMLKRKKAHEKHKGH*

Sequence 2109
Contig_0711_pos_3571_3053,
putative peptide of unknown function

55 atgttagaaacacatagattaaagctagtgaaagcctaatttgagttatacagatgaactt
tatcaattgcatacaaaataaggttagctacaaagtatacacctaaaggtattcatcagaat
aaagtagcaacccaagattttattaaaggatggatgaggcattgggatgaatatcaattt
ggttacttcatttttaattatgagagataatcacgaagtagtggggatagcgggatttgag
tatcgtaaatcatcaacaacagtttcttaatgcgtattatagaatctttccatcgat

actggtggttggttagcttttgagtcgaatggaggagattgcccgtcatttaaaaaagcat
gataccataacaccaaataattcgaacaaatcaatataatacaaatctattaaatta
gcacaaaaactcgatataattatgatgctaactgggacgatgtaattaataaaggagat
cggtgttttttaacctacaagcggttgataataactaa

5

Sequence 2110

MLETHRLKLVKPNLSYTDELYQLHTNKVATKYTPKGIHQNKVATQDFIKGWMRHWDEYQF
GYFILIMRDNHEVVGIGAGFEYRTIHQQQLNAYYRIFPSYTGVLAFESMEEIARHLKKH
DTITPKLIRTNQYNTNSIKLAQKLGYNYDANWDDVINKGDRCCFFNLQALDNN*

10

Sequence 2111

Contig_0711_pos_703_359,

putative peptide of unknown function

atggagatgatagaagaacgtaatttatcagggcttattcaaacactaactttcaatcat
cccacattcaaatcttaagagaacacattaaatcaacttaaaatactctctcattat
ttacagagcgacaccctgcaatgggtggcaattcaatcttggtcacaatgggttatgat
catgggattactgaaatccaccttgatgtaactgcacaagcgcttagatcttattacaaa
ggatattttataaaatgtcatcttaaaaaactgctcatagcggttttgacaggtggatat
tatcacgggttactagaagggttttggttaggattaacacttta

20

Sequence 2112

MEMIEERNLSGLIQTLTFNHPIIQILKENTLNQLKILSHYLPERHPAMVAIQSWSQWFTD
HGITEIHLDVDTAQAPRSYYKGIKFKHLKNTAHSVLTGGYYHGSLEGFGLGLTL*

25

Sequence 2113

Contig_0711_pos_0_334,

putative peptide of unknown function

atgttacgagttgcattagcaaagggtcggtttattaaagagttttatcgaatatttaca
caagttaatcagatagatattgcaactgtacttttaaatagacagcgacagttattgctt
acagtcgacaacattgaaatgatttttagttaaggaagcgatgtgcctacttatgtagaa
caaggtattgctgatgtaggaatagtggaagtgatattctgaatgggtcaaaaatataat
attaataaattactcgatttgccatttggttaaatgtcattttgcggttgcgcgcaagcca
gaaacatctcgctataaaaaagtagcaagTATTG

35

Sequence 2114

MLRVALAKGRLLKSFIEYLQOVNQIDIATVLLNRQRQLLLTVDNIEMILVKGSDVPTYVE
QGIADVIGVGSILNGQKYNINKLLDLPFGKCHFALAAKPETSRYKKVASIX

40

Sequence 2115

Contig_0712_pos_299_928,

is similar to (with p-value 1.0e-52)

>sp:sp|Q02499|KPYK_BACST PYRUVATE KINASE (EC 2.7.1.40) (PK).

>pir:pir|S29783|S29783 pyruvate kinase (EC 2.7.1.40) (versi

on 2) - Bacillus stearothermophilus >gp:gp|D13095|BACPK_3 B.

45 stearothermophilus phosphofructokinase and pyruvate kinase
genes. NID: g285620.

atgcctttttctaaacctataattgcgcttggtgaagtaataaccattttctctgtgaatt
agacctatagctttttcaacatatggtactaatgtttcatcaacagaatttgtaataata
actttatcagataaatctttaccttctaaatcactagcactatctgcgacaattgcatgg
50 cctacaacagatcctctaccaacaccttggcctttagcaatctcatcacctactaagtgg
attttcatcatattttagtctctttttctccagtaggtacaccagcagtaataataatt
aaatctccggttgaaactctaccagtttctactgctgttgctacagcattatttagtaaa
gcatcagttgttttacgtccttctttaacgaccggatttactcccatacaattgcacat
tgtctagcagttttttcgctaggtgttacagcaatgatctgaatgtggacgatattta
55 gaatcgtacgtgctgttgaaaccatttcagttgctgctacaatagcttttacatttaag
tttaaggcagtatgtgcaacagaaacaccaattgcatttactaatgaagtttcaactaat
ttagtacgatcacttaataattttttatag

Sequence 2116

MPFSKPIIALGEVIPFSSVIRPIAFSTYGTNVSSTEFVIIITLSDKSLPSKSLALSATIAW
 PTTDPLPTPWPLAISSPTKWIFIIFVVPFSPVGTAVIIIKSPFETLPVSTAVATALFSK
 ASVVLRLPSLTGFTPHITIAHCLAVFSLGVTAMISECGRYLEIVRAVEPLSVAATIAFTFK
 FKAVCATETPIAFTNEVSTNLVRSNNFL*

5

Sequence 2117

Contig_0712_pos_5629_0,

is similar to (with p-value 2.0e-21)

10 >gp:gp|Z99125|MLCL536_19 Mycobacterium leprae cosmid L536. N
 ID: g2398683. >gp:gp|U00013|U00013_5 Mycobacterium leprae co
 smid B1496. NID: g466868. >gp:gp|Z99125|MLCL536_19 Mycobacte
 rium leprae cosmid L536. NID: g2398683.
 atgaattttaataatttgatcaattatataagatctgtaattatggatcattacaaaaac
 cctagaaacaaagggtgtccttagacaatggctcaatgactgttgatatgaataaccctaca
 15 tgtggtgatcgcatagctttgacatttgatattgaagacggaatcattaatgatgctaag
 tttgaaggagaaggatgttcaatttcaatgtctagtgcatctatgatgactgaagcagtt
 aaaggctattcacttggtagcaatgcaaatgagccaagagtttactaaaatgatgctc
 ggtgaagactacgagattacagaagaaatggagatattgagcgcttcaagggtgtctca
 caattcccagctagaattaaatgtgcaacgcttgcatggaagcattagaaaaagggaca
 20 gtcgaaaaagaaggtaagtcagaagggtaca

Sequence 2118

MNFNNLDQLYRSVIMDHYKNPRNKGVLNDSMTVDMNNPTCGDRIRLTFDIEDGIINDAK
 FELEGCSISMSSASMMTEAVKGHSLGEAMQMSQEFKMLLEDYEITEEMGDIEALQGV
 25 QFPARIKCATLAWKALEKGTVEKEGKSEGT

Sequence 2119

Contig_0712_pos_5603_5016,

is similar to (with p-value 3.0e-27)

30 >sp:sp|P16468|MAOX_BACST MALATE OXIDOREDUCTASE (NAD) (EC 1.1
 .1.38) (MALIC ENZYME). >pir:pir|A33307|DEBSXS malate dehydro
 genase (oxaloacetate-decarboxylating) (EC 1.1.1.38) - Bacill
 us stearothermophilus >gp:gp|M19485|BACMAL_1 B.stearothermop
 hilus malic acid gene, complete cds. NID: g143164.
 35 atgcttgaactaattgatctacatcttcttctgattataaaacgtaaaaacttgacagtg
 ctggtgaagagacacctaaccatttcatcaaagggtgacgacagtgatggcctgcacgaa
 ctgcaacaccttctaggtgtatctgtagcgatataactcaagatatggtgaagaca
 atggctgatgatccaattatttttgcctatggctaatacctaactcctgaaatcaatcccaat
 gaagccaaacaggcaggtgcaaaggctcgtaggtacaggtcgatctgattttccaaaccaa
 40 attaatatgtatttagcatttccaggtatttttagaggtgcattggatggtgaagccact
 catattaacgaagatatgaaaaaggcagctgtagaagctatcggtcatttaataagacgaa
 aatgagttacatcctgattactgtataaccaggaccatttgataaaagagtagctccatca
 gtagccaaaaatgtagctaaagctgctatggaatccggtgtagcaagaattaaaattgat
 acacaagaaatatttgataaaactatgaaacttactgacttaaaatag
 45

Sequence 2120

MLELIDLHLLSYKRNKLNHVLKRHLTISSKVAHSDGLHELOHLLGVSVADILTQDMVKT
 MADDPIIFAMANPNPEINPNEAKQAGAKVVGTRSDFPNQINNVLAFPGIFRGALDVEAT
 HINEDMKKAHAVEAIVHLIDENELHPDYCIPGPFDKRVAPSVAKNVAKAAMESGVARIKID
 50 TQEIFDKTMKLTDLK*

Sequence 2121

Contig_0712_pos_2893_1970,

is similar to (with p-value 0.0e+00)

55 >sp:sp|P00512|K6PF_BACST 6-PHOSPHOFRUCTOKINASE (EC 2.7.1.11)
 (PHOSPHOFRUCTOKINASE) (PHOSPHOHEXOKINASE). >pir:pir|A27474|
 KIBSFF 6-phosphofructokinase (EC 2.7.1.11) - Bacillus stearo
 thermophilus >gp:gp|M15643|BACPFK_1 B.stearothermophilus 6-p
 hosphofructo-1-kinase gene, complete cds. NID: g143311. >gp:

gp|D13095|BACPK_2 B. stearothermophilus phosphofructokinase and pyruvate kinase genes. NID: g285620.

atgaacgctgctgttcgtgcagtgactcggacagcaatttacaataatattgaagtttat
 5 ggtgtttatcaaggttaccaaggtttacttgatgatgatattcataagcttgaattgggt
 tcagtaggggatacaattcaacgaggaggaaactttcctcttttccgcaaggtgccctcag
 ttcaaagaagaggatgtacgtaagaaagctattgagaatttacgtaagcgtggtatcgaa
 ggttttagttgttattggaggagatggcagctatagaggggcacaacgaattagtgaggaa
 tgtaaagaaattcaacaattggtattcctggtacaattgataatgatattaatggtaca
 10 gattttacaattggttttgatactgcattaaacactattattgaatcagtcgataagatt
 agagatacggcatcaagtcacgcaagaacgtttattgttgaagttatggggcgtgattgt
 ggagatttagctttatgggctggattatctgtaggtgctgaaacgattgtttaccagaa
 gtcaatacagataattaaggatgtagctgaaaagattgaacagggtattaaaaggaggaaa
 aaacattctatcggttgggtgcagaaggtgtatgagcggccaagaatgtgcagatgag
 ttaacgaagtataattaacattgatacacgagtttcagtgttaggtcacattcaacgtggc
 15 ggtagcccatctggtgctgatcgagtattagcttctcgacttgggtggatagctgttgaa
 ctattaaaacaaggcgagacagctaaaggtgttgccattaggaataatcaattaacctct
 acgccgtttgatgaaatttttgcgtgaaagtgcgcaatttaatatgcaaatgtatgaa
 ttagcaaaagaattatcaatttaa

Sequence 2122

MNAAVRAVTRTAIYNNIEVYGVYQGYQGLLDDDIHKLELGSVGDITQIRGGTFLFSARCPQ
 FKEEDVRKKAIENLRKRGIEGLVVIGDGSYRGAQRISSECKEIQTIGIPGTIDNDINGT
 DFTIGFDALNTIIESVDKIRDTASSHARTFIVEVMGRDCGLALWAGLSVGAETIVLPE
 VNTDIKDVAEKIEQGIKRGKKHSIVMVAEGCMSGQECADLTKYINIDTRVSVLGHIIQRG
 25 GSPSGADRVLASRLGGYAVELLKQGETAKGVGIRNNQLTSTPFDEIFAESDRKFNSQMYE
 LAKELSI*

Sequence 2123

Contig_0712_pos_1946_189,
 30 is similar to (with p-value 0.0e+00)
 >sp:sp|P51181|KPYK_BACLI PYRUVATE KINASE (EC 2.7.1.40) (PK).
 >pir:pir|JC4220|JC4220 pyruvate kinase (EC 2.7.1.40) - Bacillus
 licheniformis >gp:gp|D31955|BACPK2_2 Bacillus lichenif
 ormis gene for pyruvate kinase, complete cds. NID: g1041098.
 35 atgagaaagactaaaattgtatgtacaataggaccagcttcagaatcagaggaaatgctc
 gaaaaactaatgaatgcaggaatgaacgttgcgcgtttaaatttctcacatggtatgcat
 gaagaacataaagcaagaattgatacaattcgtaaagttgctaaacgttttaataaaaaca
 attggttgtttattggatactaaagggccagaaattcgtacgcacaatatgaaagtgtga
 cttattgttttagaaaaaggaagaagtcattgtcagtatgaatgaagttgaaggaaca
 40 cctgaaaaattctctgtaacatatgaaaatctaataatgatgtcaatattggatcatat
 atactattagatgatggttttagttgaacttcaagtcacaagaaattaacaaagataaaggc
 gaagttaaattgtgatatttaataactggtgaattaaaaataaaaaagggtttaactta
 cctggtgttaaagttaatttacctggtatcactgataaagatgccgatgatcagattt
 ggtataaaggaaaaatgtagactttatagctgcaagttttgtaagacgtccaagtgtggt
 45 ttagatatccgtcaaattcttgaagaagaaaaagcagaaataacaattttccctaaaatc
 gaaaaccaagaaggtatcgataatattgaagaatttcttgaagtatctgatggattaatg
 gtacacgtggtgatattgggtgttgaaattccaccagaaagcgtaccaatggttcaaaaa
 gatttaattagaaaatgtaataaattaggaacacctgtaattactgcgactcaaattgctt
 gattctatgcaacgtaataccacgtgcgacacgtgcagaagcaagtgcgtagctaatagca
 50 atatacgtatggtactgacgtgtaatgttatcaggcgaaaactgcagcaggtcaatatcct
 gaagaagctgttaaaactatgcgtaataattgcagtttctgctgaagcagcgcaagactat
 aaaaaattattaaagtgcgtactaaattagttgaaacttcattagtaaatgcaattggt
 gttctgttgacacactgccttaaaacttaaatgtaaaagctattgtagcagcaactgaa
 agtgggttcaacagcacgtacgatttctaataatcgctccacattcagatatattgctgta
 55 acacctagcgaaaaaactgctagacaattgtgcaattgtatggggagttaaatccggtcggt
 aaagaaggacgtaaaacaactgatgctttactaaataatgctgtagcaacagcagtagaa
 actggtagagtttcaaacggagatttaattattattactgctggtgtacctactggagaa
 aaaggaaatatacaaatatgatgaaaatccacttagtaggtgatgagattgctaaaggccaa
 ggtgttggttagaggatctgtttaggccatgcaattgtcgcagatagtgctagtgattta

gaaggtaaagatttatctgataaagttattattacaaattctgttgatgaaacattagta
ccatatgttgaaaaagctataggtctaattacagaagaaaatgggtattacttcaccaagc
gcaattataggttttagaaaaaggcatacctactgttggtgtagaacaagcaactaaa
gaaattaaaaatgatatgttagtgacttttagatgcgtcacaaggtaaagtgtttgaaggt
5 tatgctaacgtcctttaa

Sequence 2124

MRKTKIVCTIGPASEEEMLEKLMNAGMNVARLNFSGSHEEHKARIDTIRKVAKRLNKT
IGLLDITKGPEIRTHNMKDGILIVLEKGKEVIVSMNEVEGTPEKFSVTYENLINDVNI
10 ILLDUGLVLEQVKEINKDKGEVKCDILNTGELKNKKGVNLPGVKNLPGITDKDADIRF
GIKENVDFIAASFVRRPSVLDIRQILEEEKAETIFPKIENQEGIDNIEEILEVSDGLM
VARGDMGVEIPPESVPMVQKDLIRKCNKLGKPVITATQMLDSMQRNPRATRAEASDVANA
IYDGTDAVMSGETAAGQYPEEAVKTMRNIAVSAEAAQDYKKLLSDRTKLVETSLVNAIG
VSVAHATLNLNVKAIVAATESGSTARTISKYRPHSDIIAVTPSEKTARQCAIVWGVNPVV
15 KEGRKTTDALLNNAVATAVETGRVSNGLIIITAGVPTGEKGTNMMKIHVLVGDEIAKGQ
GVGRGSVVGHAIVADSASDLEKDLSDKVIITNSVDETLVPYVEKAIGLITEENGITSPS
AIIGLEKGIPTVVGVGEQATKEIKNDMLVTLASQGVFEQYANVL*

Sequence 2125

20 Contig_0714_pos_181_783,
is similar to (with p-value 2.0e-33)
>gp:gp|D50453|D50453_106 Bacillus subtilis DNA for 25-36 deg
ree region containing the amyE-srfA region, complete cds. NI
D: g1805369.
25 atgcttccaaagctgtttacgttggttaaagatttatataaaaaggttaagtcataccttct
ttcgtaaactcgatctgctgtagcaggtattttaacaactgggtcatgcgcaccttattg
tttttagctgtactaggtgttggtgtaactggcggttacgcttagttcagaaaaatccacca
gcatcagttttccaacatgcatttaggtcctataggtaaaaatatttttggcgttagtaata
tttgacagcagcaatgtcctcagtaattggttctgcatatacaagcgcaacatttttaaaa
30 acactacacaaatcggtactcaataaaaaataatcttatcggtattacatttattggaatt
tcaacttttgttttcttatttattggttaaaccggtgagtttacttataatagctggtgcg
attaatggttggtttcttaccatcacattaggtgcaattctcattgcaagtaggaaaaaa
tctatcggttggaattaccaacacccaacatggatgcttggttttgggtattatagccgta
attgtcacataatgactggtatcttttcattacaagatttagcaagtccttggaaggt
35 taa

Sequence 2126

MLPKLFLVVKIYIKGKSYLPFVNRSVAVAGILTTGVMRTLLFLAVLGVVVTGVTLSSENPP
ASVFQHALGPIGKNIFGVVIFAAAMSSVIGSAYTSATFLKTLHKSLNKNLIVITFIVI
40 STFVFLFIGKPVSLIIAGAINGWILPITLGAILIASRKKSIVGNYQHPTWMLVFGIIAV
IVTIMTGIFSLQDLASLWKG*

Sequence 2127

45 Contig_0714_pos_1712_2215,
is similar to (with p-value 4.0e-46)
>sp:sp|P54452|YQEG_BACSU HYPOTHETICAL 20.1 KD PROTEIN IN NUC
B-AROD INTERGENIC REGION. >gp:gp|D84432|BACJH642_91 Bacillus
subtilis DNA, 283 Kb region containing skin element. NID: g
2627053. >gp:gp|Z99117|BSUB0014_48 Bacillus subtilis complet
e genome (section 14 of 21): from 2599451 to 2812870. NID: g
50 2634966.
atgccaaatgcatatgtgaaatcaatatttgaaattgatatagaaaaacttgccgatagt
ggtgttaaaggtatcataactgatttagataatacacttggtggttggtggtggtggtggtg
cctactaaggggtgttaaatcatgggttgctaaaggttaagatttaggaataactgtcaca
55 attgtgtcaataataataaaaagtcgagtatcaagtttctcaagtaatttaggtgtagat
tatatattcaaagcacgtaaaccgatggggaagcctttaagatggctattaaaaaaatg
aaaattcaaccgagagaaaccggtgttgtaggagatcaaatgcttactgatgtgttggt
ggcaattgtaattggttatatacaattatggttagtacctgttaaacggactgatggatta
attacaaagtttaatcgattaattgaaagacgattattaaatcatttttagaaaaaaaggt

tatattaaatgggaggaaaattga

Sequence 2128

MPNAYVKSIFEIDIEKLADSGVGKIITDLNNTLVGWDVKEPTKGVKSWFAKAKDLGITVT
 5 IVSNNNKSRSVSFSSNLGVDYIFKARKPMGKAFKMAIKMKIQPRETVVVGQMLTDVFG
 GNCNGLYTIMVVPVKRTDGLITKFNRLIERRLLNHFRRKGYIKWEEN*

Sequence 2129

Contig_0714_pos_2822_3316,
 10 is similar to (with p-value 6.0e-39)
 >sp:sp|P54453|YQEH_BACSU HYPOTHETICAL 41.0 KD PROTEIN IN NUC
 B-AROD INTERGENIC REGION. >gp:gp|D84432|BACJH642_92 Bacillus
 subtilis DNA, 283 Kb region containing skin element. NID: g
 2627063. >gp:gp|Z99117|BSUB0014_47 Bacillus subtilis complet
 15 e genome (section 14 of 21): from 2599451 to 2812870. NID: g
 2634966.
 atgatagatattccattagacgaaaaatcatttatgtttgatacaccaggtatcattcaa
 tcacatcaaatgacaaattatgtatatgaaaatgagttgaaaatcattatacctaaaaat
 gaaataaagcaacgtgtgtatcaacttaatagaaaaacagacattatcttcggaggattg
 gcacgcattgattatgtatctgttggtataaagaccactgtttgtttctttcaaatgat
 20 ttaattatcatagaactaaaaccgagaaagctaattgatttatggaaatccaattaggc
 gcattgctttcaccgcctcaagatgcacaacaatttaattcttaattgatgtaaaagcagta
 agactggaaactggtaaaactaaacgtgacatcatgatattctggttaggattcataact
 attgatgctggtgcaaaagtgatagttcgtgttccaaaacatgtagatgttattttaaga
 25 aattcaattctttaa

Sequence 2130

MIDIPLDEKSFMDTPGIIQSHQMTNYVYENELKIIIPKNEIKQRVYQLNEKQTLFFGGL
 ARIDYVSGGKRPLVCFSSNDLNIHRTKTEKANDLWKSQLGALLSPQDAQQFNLNDVKAV
 30 RLETGKTKRDIMISGLGFITIDAGAKVIVRVPKHVDVILRNSIL*

Sequence 2131

Contig_0714_pos_6740_7852,
 is similar to (with p-value 0.0e+00)
 35 >pir:pir|A55856|A55856 llm protein - Staphylococcus aureus
 gtgaggtacaacttattcaatgaaggtgaactgatgtatacactattacttatagctttt
 actatgatagtcagtttaataattacaccattattattgtaatatcaaaaaaattagat
 ttagtagatcgtcctaatttcagaaaagtacatacgaacctaattcagtgatgggagga
 agcgtcatttttattttctttcttaataagggatttggtcggacaccctattgaacgtgag
 40 gtaaacccgcttatattaggtgcaattacaatgtatatggttggtattgattgatgaatt
 tacgactcaagacctattttaaagtttagcaggtcaaattggtgcagctttaattgttacg
 ttttatggaattacaatagactttatttcattgccaaattggtccaacgattcattttggc
 atattcagcattcctattacagtaatatggattgtagcaattaccaatgctattaatcct
 atcgacggacttgatggacttgctcaggcgtctcagcattggcattaatgactattgga
 45 ttcacgcgtattttacaagcgaacatatattatcatgatttgctgtgacttttaggg
 tctttacttggtttcttattctataactttcaccagcgaaaattttcctaggtgtagt
 ggtgcattaatgataggatttattatcggtttcttactcctactcggctttaagaatatac
 acatttattgcattattctttcctatagttatattagcgggtgccatttattgatacatta
 tttgcaatgattcgtcgaatgaaaaagggaacatatataatgcaagcggacaagtcacat
 50 ttacatcataaattacttgcttttaggatatacgcatagacaaaccgttttacttatttat
 tcaatagcgattatgtttagtttatctagtgttatcctctattttatcccaaccgttgggt
 gcacttatgatgttcattctcattgtctttacgattgagttgatcggtgaatttactgga
 ttaatagatgataattatcgaccaatattaaatttaattacaaaaaaagggaatggtgaag
 caacatcattatgatgagcatcaccggttcataa
 55

Sequence 2132

VRYNLFNEGELMYTLLLIAFTMIVSLIITPIIIVISKKLDLVDPRNFRKVHTKPISVMGG
 TVILFSFLIGIWLGHPIEREVKPLILGAITMYMVGILIDDIYDLRPLYKLQIVAAALIVT
 FYGITIDFISLPIGPTIHFIFISIPITVIWIVAITNAINLIDGLDGLASGVSAALMTIG

FIAILQANIFIIMICCVLLGSLGFLFYNEHPAKIFLGDSGALMIGFIIGFLSLLGFKNI
TFIALFFPIVILAVPFIDTLFAMIRRMKKQGHIMQADKSHLHHKLLALGYTHRQTVLLIY
SIAIMFSLSSVILYSQPLGALMMFILIVFTIELIVEFTGLIDDNYRPILNLITKKGNK
QHYYDEHHRS*

5

Sequence 2133

Contig_0714_pos_8431_7913,

is similar to (with p-value 7.0e-41)

>gp:gp|Z99122|BSUB0019_48 *Bacillus subtilis* complete genome
(section 19 of 21): from 3597091 to 3809700. NID: g2636029.
>gp:gp|U56901|BSU56901_2 *Bacillus subtilis* putative transcrip-
tional regulator (yvhJ), Ycr59c/YigZ homolog (yvhK), histid-
ine kinase (degS), transcriptional regulator of degradation e-
nzyme (degU), (degV), (comFA), (comFB), (comFC), flagellar p-
rotein (yviB), negative regulator of flagellin (flgM), flage-
llar protein (yviC), flagellar-hook associated protein 1 (flg-
K), flagellar-hook associated protein 3 (flgL), (yviE), tra-
nsmembrane protein (yviF), (csrA), flagellin (hag), flagella-
r protein (yviH), flagellar hook-associated protein 2 (fliD)
20 , flagellar protein (fliS), flagellar protein (fliT), sigma-
54 modulator homolog (yviI), and (secA) genes, complete cds.
NID: g1762326.

gtgatccggttcgatataaagatgcaactcataattgttcagcctatactgtcggacca
gagatgaatattcaaaaggcaaacgacgatggcgaaccaagtggaacagctggcatccca
25 atgcttgaaatactgaaaaaacaagagatacacaatgtttgtgctcgtcgtgacacgctac
ttcgggtggtatcaagtttaggtgcaggcggtcttattagacatatagcggcgccgtgcgt
gatgtgatatatgatataagtagagtcgaactaagagaagctattccagtaaccgttacg
ttagattatgatcagacaggtaaatttgaatatgaacttgccctactacattcttatta
agagaacaattttataccgataaagtaagttatcaaattgacgtagtaaaaaatgaatat
30 gatgctttttatagacttttttaaatcgaattacttctggaaattatgatttgaacaagaa
gaccttaactattaccttttgatattgaaaccaattaa

Sequence 2134

VIRSIHKDATHNCSAYTVGPEMNIQKANDDGEPSGTAGIPMLEILKKQEIHNVCVVTRY
35 FGGIKLGAGGLIRAYSGAVRDVIYDIGRVELREAIPTVTLDYDQTKFEYELASTTFL
REQFYTDKVSQIDVVKNEYDAFIDFLNRITSGNYDLKQEDLKLPPFDIETN*

Sequence 2135

Contig_0714_pos_6433_5357,

40 is similar to (with p-value 2.0e-25)

>sp:sp|P54595|YHCK_BACSU HYPOTHETICAL 40.7 KD PROTEIN IN CSP
B-GLPP INTERGENIC REGION. >gp:gp|X96983|BS75DGREG_12 *B. subtilis*
chromosomal DNA (region 75 degrees: cspB upstream of glp
PFKD operon). NID: g1239975. >gp:gp|Z99108|BSUB0005_180 *Baci-*
45 *llus subtilis* complete genome (section 5 of 21): from 802821
to 1011250. NID: g2633055.

atggaaatgtttgaagctatcatatataacatatctgtcatggtggcaggtatatattta
tttcataggttacaatattctgaaaataaaagaatgattttttctaaagaatatgtaaca
gtactaatgacattcgtttctttacttttagcggcataccctatcccatctcaaaacgaa
50 tacctcgtccatttaacattttgtacctctttgttttaggacgttataccaacatgata
tatacactcacggctgcttttatcgtatcttttagtcgatgtatttatctttggaaactca
attatttttggtattacattaatcgttattgcaggtattgtcagtgagtgaggaccattc
ttaagcaaacgatacattctcttacttatttttaatttgattagcattatcattttg
ttatttttagcattattaagccctatttatgaactcgtagagatttttagtgcttatccct
55 atttcattttattattacaattgcttcagcaataacattcgttgatataatggcactttttc
tcttttagtcaatcgttatgaaaatgaagataaacgattatcttacaggtctaggtaat
gtgaaagaatttgatagacacttaaatgaggtctcaagtaagctgaagaaaagaaacaa
agtttagccttacttctcattgatattgatggctttaaagatgtaaacgatcattattca
accaatcaggagatgctgttctcaacaaatgtctcaactattaaaaaactatgtccca

Sequence 2136
MEMFEAIIYNISVMVAGIYLFHRLQYSENKRMIFSKEYVTVLMTFVSLLLAAYPIPFQNE
YLVLHTFVPLLFLGRYTNIYTLTAAFIVSLVDVFI FGNSIIYGITLIVIAGIVSAVGPF
LKQNDIISLLILNLISIIILLFLALLSPIYELVEILVLPISFIITIASAITFVDIWHFF
SLVNRAYLENEDKYDLYLTGLGNVKEFDRHLNVESSKAEKKQSLALLIDIGDFKDVDNDHYS
HQSGDAVLKQMSQDLLKNYVPNQFKIFRNGGEEFSVIRDTYLDQSVKLAEINIRSGVEKSS
FHLPNKEVIKLSVSGVGYLTQEDRKSQRKVKFDADDMVHVAKSEGRNKVMFNPIVKL*

Sequence 2138
MNRIAHSYGLHDTYSFVTSTAIIFSLNDRSTRLIRIRERTTDLEKIALTNSLSRKISSK
QLTIDEAKSELLQLKRASLQYSFLTNLIAAFVACGFFLFMEGGVASDAWIACLAGGIAFL
TFSFVHFSI*

Sequence 2140
MIPGVSNIINFSSNGISIIISKVVPGKRVDVTTSFSP¹TLCSINLLINVD²FPTFV³VPTMYTS
SLFLTWFIDCNNSSIPQPLFAEIRTTSSASNPYLRA⁴DFLNHSFTR⁵*

532

cttgtatatgcagaaccaattactgaggacattgctgctgcaaataattactacgcaaaa
 atatTTTTTtacctataggacctaataatgcatgttggaactgatgctggtggattttctgaa
 ctaagcgtaacgccagttacaacaacacctagtagcagctaaaaacaataaggtgcatg
 acaccagttgttaaaatacctgctacagcagatcgatttacgaaaggaaggtatgactta
 5 ccttttatataa

Sequence 2142

MRIAPNVIGRIQPLIAPAIISKLTGLPINKKTKVEITINVITIRLFLSNDLCSVFKNVA
 LVYAEPITEDIAANITTPKIFLPIGPACWKT DAGGFSELSVTPVTTTPSTAKNNKVRM
 10 TPVVKIPATADRFKGRYDLPI*

Sequence 2143

Contig_0715_pos_1038_1508,

is similar to (with p-value 4.0e-35)

15 >sp:sp|Q02134|HIS7_LACLA IMIDAZOLEGLYCEROL-PHOSPHATE DEHYDRA
 TASE (EC 4.2.1.19) (IGPD). >pir:pir|G45734|G45734 HisB - Lac
 tococcus lactis subsp. lactis >gp:gp|U92974|LLU92974_6 Lacto
 coccus lactis unknown gene, partial cds, and HisC (hisC), un
 known, HisG (hisG), unknown, HisB (hisB), unknown, HisH (his
 20 h), HisA (hisA), HisF (hisF), HisIE (hisIE), unknown, unknow
 n, LeuA (leuA), LeuB (leuB), LeuC (leuC), LeuD (leuD), unknow
 wn, IlvD (ilvD), IlvB (ilvB), IlvN, IlvC (ilvC), IlvA (ilvA)
 , AldB (aldB) and aldR (aldR) genes, complete cds. NID: g256
 5137.
 25 atgttaacgctatttacttttcatagtggattaactttatctattgagggcactggagat
 acgtatgttgatgatcatcatataactgaagatataggtatagttattggacaattactt
 cttgaattaataaagactcaacaaagttttacaagatatggtgctcatatgtacccatg
 gatgagggcgttgcgcgaacagtagtggaacattagtggtcgtccatatttctcattta
 agcaagttgagcgtcaaaaggttaggaacttttgacactgaactagttgaagaattttt
 30 agagcattgataattaatgcgcgattaaccgttcacattgacttattaagaggtggaaat
 acacatcatgagattgaggcaatatttaaatcttttgcaagagcattaaagatttctct
 gcacaaaatgaagatggacgtattccatcgtctaaaggagtaattgaatga

Sequence 2144

35 MLTLFTFHSGLTSLIEATGDTYVDDHHITEDIGIVIGQLLLELIKQQSFTRYGCSYVPM
 DEALARTVVDISGRPYFSFNLSAQKVGTFDELVEEFFRALIINARLTVHIDLRRGN
 THHEIEAIFKSFARALKISLAQNEDGRIPSSKGVIE*

Sequence 2145

40 Contig_0715_pos_1607_1990,

is similar to (with p-value 1.0e-21)

>sp:sp|Q02132|HIS5_LACLA AMIDOTRANSFERASE HISH (EC 2.4.2.-).
 >pir:pir|I45734|I45734 HisH - Lactococcus lactis subsp. lac
 45 tis >gp:gp|U92974|LLU92974_8 Lactococcus lactis unknown gene
 , partial cds, and HisC (hisC), unknown, HisG (hisG), unknow
 n, HisB (hisB), unknown, HisH (hish), HisA (hisA), HisF (his
 F), HisIE (hisIE), unknown, unknown, LeuA (leuA), LeuB (leuB)
 , LeuC (leuC), LeuD (leuD), unknown, IlvD (ilvD), IlvB (ilv
 B), IlvN, IlvC (ilvC), IlvA (ilvA), AldB (aldB) and aldR (al
 50 dR) genes, complete cds. NID: g2565137.
 gtgcaaaaagctgaagctatcgtacttccaggtgttggaacattttcaggatgcatgcat
 tctatagaagaaaaaagcattaaagatatgcttaaaaatatacatgataaacgataatt
 ggaatattgttttaggtatgcaattactttttcaacatagcgcagaaggtgacgttagtgga
 ttggaacttgcgccgggaatatagtgccaatccaatcatctcatctcattcctcat
 55 ggttggatgaattaaagagtacacatcccttactgcaaagtgtgtgtattttgttcat
 tcatatcaagcagaatgtcagaatatgtagtagcttatgtgactatggtacaaagatt
 ccgggagtcattcaataccgatga

Sequence 2146

VQKAEATVLPVGVGHFQDAMHSIEEKSIKMLKNIHDKPIIGICLGMQLLFQHS AEGDVSG
LELVPGNIVPIQSSHPHPLGWNELKSTHPLLQSDVYFVHSYQAEMSEYV VAYADYGTKI
PGVIQYR*

5 Sequence 2147
Contig_0715_pos_2067_2615,
is similar to (with p-value 9.0e-31)
>sp:sp|P54471|YQFN_BACSU HYPOTHETICAL 23.7 KD PROTEIN IN CCC
A-SODA INTERGENIC REGION. >gp:gp|D84432|BACJH642_139 Bacillu
10 s subtilis DNA, 283 Kb region containing skin element. NID:
g2627063. >gp:gp|Z99116|BSUB0013_228 Bacillus subtilis compl
ete genome (section 13 of 21): from 2395261 to 2613730. NID:
g2634723.
gtgattcaaggaccttataaggctgctaaaaagaaatattgcaaattatgaattaaatcaa
15 cagggtgatgtacgtctagcgatgggtctaagcgttataaacacagaagaccaaattgat
aatataactggttgtggtatgggagggccattaattgcaaaaatattaaacgatggaaaa
gataaattagtttaaccatccaagactcatactacaaagcaacatacaaaactcaagcatta
agacaaactcttaataaaactttcatatgaaatcgttgatgaaagaatcattgaggaaaag
ggtcacatatatgaaatcgtggttagctgagtttaataaacttagttaaattaaatata
20 ttacaagaaaaattcggaccatttttacttagagaatgtaataacatttttcaaaaaaaa
tggcaaagagagtttagaagcactgcgtgatataaaatccaattgaattcaacatcacat
catgagagactaaaagaaatagaagatgaaattaacttaatacaagaggtgtaattaat
gaaaattag

25 Sequence 2148
VIQGPYKAAKRNIANYELNQVDVRLGDGLSVINTEDQIDNITVCGMGGPLIAKII,NDGK
DKLVFHPKILIQSNIQTQALRQTLNKLSEIVDERIIIEKGHIYEIVVAEFNNNLV,KLNI
LQEKFGPFLRECNNIFQKKWQRELEALRDIKSQLNSTSHHERLKEIEDEINLIQEVLIN
EN*

30 Sequence 2149
Contig_0715_pos_2623_3705,
is similar to (with p-value 2.0e-54)
>sp:sp|P53434|YRP2_LISMO HYPOTHETICAL 41.4 KD PROTEIN IN RPO
35 D 3'REGION (ORFA2). >gp:gp|U17284|LMU17284_3 Listeria monocy
togenes major sigma factor (rpoD) gene, partial cds, and dow
nstream orfA1 and orfA2 genes, complete cds. NID: g687597.
atggaagtttttaataatcacgtttccattcatcaagctgaatcatgggataatgtagga
ttattaattggtaatgataaagttagatattacaggtattctgacaacactcgactgcacc
40 gatgatggttgttaaccaagcaattgaacttaataccaataccatcattgctcatcatcca
cttattttcaaaggagtaaaacgtatcgttgaagatggatatggtagtataattcgtaaa
cttatccaaaataatatcaatcttatagcattacacactaatcttgatgtaaatcctaaa
ggtgtcaatcgaatgttagcggatcaaataaggttttagagaacatatcaatgattaatata
aatagctcatattattacaaagttcaaacttttatacctaaaaattatattgaagatttc
45 aaagacagttttaaacgaacttgatttagctaaagaaggtaattacgaatattgtttcttt
gaaagtgaaggtaaaaggcaattttaaacagtaggtgatgcaagtccttatataggggaag
ttagatagttatcgaatatgttgatgaaataaaaacttgagtttatgataaaagacaatgaa
ttagaataaactaaacgtgctatttttagataatcaccatacgaacaccagtttttgat
tttattaaaaatgaacaaagaaagtgatgatttagggattattggacaattaaaccaa
50 actatgacttttagatgaattttctgaatatgccaaaaaacagctcaatataccgagcgta
cgatatacaggtcaacatgatagtcgaattaaagaaagtagctatcataggtggttcagggt
ataggatttagatataaagctagccaacttgagcagatgtttttgtagtgatatt
aaacacccatgatgcttttagatgctaaaatccaaaatgtaatttattagacatcaatcat
tatagttagtattgttatgaaagaaggattaaaagaattattagaaaaatggttattttaa
55 tatgaaaatcaattttcaatatatgcttctgaaatcaacacagatccattttaaataaaa
taa

Sequence 2150
MEVLNNHVPFHFQAESWDNVGLLIGNDKLDITGILTTLDCTDDVVNQAIELNTNTIIAHP

LIFKGVKRIVEDGYGSIIRKLIQNNINLIALHTNLDVNPKGVRMLADQIGLENISMINT
 NSSYYYYVQTFIPKNYIEDFKDSLNLGLAKEGNYEYCFESEGKGQFKPVGDA SPYIGK
 LDSIEYVDEIKLEFMIKDNELEITKRAILDNHPYETPVDFIKMNKESEYGLGIIGQLNQ
 TMTLDEFSEYAKQLNIPSVRYTGQHDSPIKKVAIIGSGIGFEYKASQLGADVFTGDI
 5 KHHDALDAKIQNVNLLDINHSEYVMKEGLKELLEKWLFKYENQFPIYASEINTDPFKYK
 *

Sequence 2151

Contig_0715_pos_3740_4072,

10 putative peptide of unknown function

atgtcaaacatccatttgaacactttaatttagatgagaatttaattgaagctgttaaa
 aatctcaattttgaaaaaccgactgaaatccaaaatagaatcataccgagaattcttaaa
 ggaacaaatttaataggacaatctcaactggaactggaagtcacacgcttttcttta
 atattgatggagatcgttcttttctgaagcagtacaattggacatacaagtacctggtg
 15 cttttctcaatttggaaactacagatggaaacgattacgaaaaatgttgtcattgaaaaat
 gttattttggtgttcagatcatcctaaaatga

Sequence 2152

MSKHPEHFNLNLEAVKNLNFEPTEIQNRIIPRILKGTNLIGQSQTGTGKSHAFLL
 20 ILMEIVLFLKQYNWYTKYLVFLNLELQMERLRKMLSLKNVILAVQIILK*

Sequence 2153

Contig_0717_pos_4083_5096,

putative peptide of unknown function

25 atggaacgattttgtgtgtaaatcaaatatactatattcaaatgaatccgttagaagcc
 aaattttaaaccgagcgtcttaagatcatggaaaactgatcaggcagatgctcataagctt
 gcttgttaggaccgacgcttaaacaaacagacaacttacctatacatgagttaatattc
 tttgaattaagagaacgcgtccgttttcatctagaaatcgagaatgaacaaaatcgactt
 aaatttcagatccttgaattactccatcaaacattccctggtttagaaagattatttagt
 30 agtcgatattcaatcattgcactcaacatcgagaaatctttactcatccagacatggtt
 ctgtgatcgcagaaggagggtactgattacacatatattcaattctacagataagggaatg
 tcaatggataaagctacaaaatattgcacttcaattaagggtgattgctcaagaaagctat
 cctaattgtcgatagacattcctttctagtgcgaaaaattacgcttacttattcaacaatta
 aaacaatctattcatcatatcaacaatttagatgatgccatgattcaattagcacaacaa
 35 ctcgattattttgaaaatattcattcgatacctggatttggaagctaagcacagctatg
 attattgggaagattgggtgatattaagcgatttaaatcaataaacaactcaatgctttt
 gttggcattgatatacaacgatataatcaggtcatcacactgtagagataccatcaac
 aagcctgttaataaaaaagcgagaaaaactttttttgggtgattatgaatataaaga
 ggcagcatcattatgacaatcatgtcgatgattattactacaaactaagaaagcgcct
 40 aatgagaaacctcataagactgccatcattgctgtataaatcgattattaaaaacaatt
 cattatcttgtaattgaattataaattgtacgattatcaaatgtcaccacattag

Sequence 2154

45 MERFCCVNQINYIQMNPLEAKFKTSALRSWKTDQADAHKLACLGPTLKQTONLPIHELIF
 FELRERVRFHLEIENEQNRLKFQILELLHQTFPGLERLFSSRYSI IALNIAEIFTHPDMV
 LDIDKEVLITHIFNSTDKGMSMDKATKYALQLRVIAQESYPNVDRHSFLVEKLRLLIQQL
 KQSIHHIKQLDDAMIQLAQQLDYFENIHSIPGIGKLSTAMIIGKIGDIKREKSNKQLNAF
 VGIDIKRYQSGHTHCRDTINKRGNKARKLLFWVIMNIIRGQHHDYDNHVVDYKYLRKQP
 NEKPHKTAIIACINRLKTIHYLVNMYKLYDYQMSPH*

50

Sequence 2155

Contig_0717_pos_8693_0,

is similar to (with p-value 1.0e-49)

55 >sp:sp|P42423|YXDL_BACSU HYPOTHETICAL ABC TRANSPORTER ATP-BI
 NDING PROTEIN IN IDH 3'REGION. >gp:gp|D14399|BACIOLO_13 Baci
 llus subtilis 15 kb chromosome segment contains the iol oper
 on. NID: g709980. >gp:gp|Z99124|BSUB0021_68 Bacillus subtili
 s complete genome (section 21 of 21): from 3999281 to 421481
 4. NID: g2636442. >gp:gp|D45912|D45912_2 Bacillus subtilis g

enome sequence between the *iol* and *hut* operon, partial and complete cds. NID: gl408482.

atgggtccttctggatcaggtaaaacgactttactcaatgtgttaagttcaatagatact
 atttcagaaggaactgtggaagttgaaggcaaaagaattaataaaactgagccacaaagaa
 5 gtggcaaattttcgaaaacaacatctcggttttatttttcaagattatagcgttttaccc
 acattaacagtaaaagaaaatattatgctaccactctcagtacaaaaattccataaatat
 gaaatggaacaaaattataaagaagtggtgaggttaggtatttataacctgggaaat
 aaatatccaagtgaattttctggcgtcagcaacaacgtacggcggcagcccggtcattc
 gtrcataaaccaacgattattttcgcagatgaacctactggcgcattagattctaaaagt
 10 gctcaagatttgttacaccgtctagaagatatgaataaacaatttaattcaaccattatg
 atggtgacacatgatccttcagccgctagttacgctgagagagtcattatgttgaaagac
 ggtgatatacactcagaaatctaccagggtaacgattcaaaacaaacattttaccaagaa
 attatgaaacttcc

15 Sequence 2156

MGPSGSGKTTLLNLVSSIDTISEGTVEVEGKEINKLSHKEVANFRKQHLGFIQDYSVLP
 TLTVKENIMLPLSVQKFHKYEMEQNYKEVAEALGIYNLGNKYPSEISGGQQRTAAARAF
 VHKPTIIFADEPTGALDSKSAQDLLHRLDMNKQFNSTIMMVTHTDPSAASYAERVIMLKD
 GDIHSEIYQGNDSKQTFYQEIMKLP

20

Sequence 2157

Contig_0717_pos_8309_8007,

putative peptide of unknown function

atggagaaaaaggagtattaatatgaaaaaagtatttatgatcataagtatacttaccata
 25 actgttactttaagtgcattgtggaggttctggaaaacaaaaagagccatctaaggaaagt
 caaaaatctgataaatatgattatgtttattatgaatatataaatgatggagattctgaa
 acgcaaatgttgagattaaatataaagataaaaaaggtaaatcacatatagaaaaagct
 gatttagatcacgtgtatgaacataactaggtgatggtaataaaaaacatatatgata
 tga

30

Sequence 2158

MEKRSINMKKVFMIIISILTITVTLACGGSGKQKEPSKESQKSDKYDYVYYEILNDGDSE
 TPNVEIKYKDKKKGKSHIEKADLDHVYEHILGDGNKKPYMI*

35 Sequence 2159

Contig_0717_pos_3267_2203,

is similar to (with p-value 0.0e+00)

>sp:sp|P38021|OAT_BACSU ORNITHINE AMINOTRANSFERASE (EC 2.6.1
 .13) (ORNITHINE--OXO-ACID AMINOTRANSFERASE). >pir:pir|B53370
 40 |B53370 ornithine aminotransferase - Bacillus subtilis >gp:g
 p|D78193|BACGNTZA_27 Bacillus subtilis 36kb sequence between
 gntZ and trnY genes encoding 34 ORFs. NID: gl064780. >gp:gp
 |X81802|BSROCEDEF_1 B.subtilis rocD, rocE and rocF genes. NID
 : g550310. >gp:gp|Z99124|BSUB0021_139 Bacillus subtilis comp
 45 lete genome (section 21 of 21): from 3999281 to 4214814. NID
 : g2636442.

atggatatgctttcggcctactcggcagtgaaatcaaggatcatcgacaccaagaattatt
 caagcattgaaagatcaagcagataaagtcacttttagtatcacgtgcttttcatagtgat
 aatcgggtcaatggtatgagaaaatattgaaactcgcaggtaaagacaaagcattgcct
 50 atgaatacgggagcagaggcggttgaaacagctttaaagctgctcgctggtgggcttat
 gatgttaagggtattgagccgaacaaagctgaaattatcgcttttaattggtaatttccat
 ggacgtacgatggcaccagtatcattgtcttcagaagctgagtatcaacgaggctatggt
 ccattgttagatggcttttcgaaaagttgagtttggtgacgttaattcaattaaaagcagca
 attataaaaaatacagcagcaatttttagtagaacctatacagggagaagcagggttaac
 55 gtaccaccagaaggatatttgaaaacaattagagaattatgtgatgaacatcaaatttta
 ttattgtgatgaaattcaagcaggatttaggacgttcaggaaaattatttgcaacggat
 tgggatcatgtaaaaccggatgtttatatttttaggaaaagcgttaggtggaggggtattt
 cctatctcggtagttcttgagataatgaggtattagatgtatttactcctggctcacat
 ggttctacatttgggtgaaatccactagcgagtgacgtttctattgcagctatagatgtc

attcaattgacgaggattttacctggtcggttcattagaattaggagaatattttaagt:tgaa
 ttgaaaaaaatttgagcatccatctatttaaagaagttaggggacgaggattatttatcggt
 attgaattacatgaaagtgcagaccatattgtgaagctttgaaagaacaaggattatta
 tgtaaagaaactcacgacaccgttattagatttgacacctcattagtgataacgaaagaa
 5 gagttagacatggcttttagagaagattaagagtgtatttgcatag

Sequence 2160

MDMLSAYSAVNQGHRHPRIIQALKDQADKVTLSRAFHSNGLGQWYEKICKLAGKDKALP
 MNTGAEAVETALKAARRWAYDVKGIEPNKAEIIAFNGNFHGRMAPVSLSSAEYQRGYG
 10 PLLDGFRRKVEFGDVNQLKAAINKNTAAILVEPIQGEAGINVPPEGYLKTIRELCDEHQIL
 FIADEIQAGLGRSGKLFATDWDHVKPDVYILGKALGGGVFPISVVLDNEVLDFVTPGSH
 GSTFGGNPLASAVSIAAIDVIDEDLPGRSLELGEYFKSELKKIEHPSIKEVRGRGLFIG
 IELHESARPYCEALKEQGLLCKETHDVTIRFAPPLVITKEELDMALEKIKSVFA*

15 Sequence 2161

Contig_0717_pos_1974_850,

is similar to (with p-value 0.0e+00)

>sp:sp|P50735|YPCA_BACSU HYPOTHETICAL 46.7 KD OXIDOREDUCTASE
 IN RECQ-CMK INTERGENIC REGION (EC 1.4.1.-). >gp:gp|Z99115|B

20 SUP0012_236 Bacillus subtilis complete genome (section 12 of
 21): from 2195541 to 2409220. NID: g2634478. >gp:gp|Z99116|

BSUB0013_8 Bacillus subtilis complete genome (section 13 of
 21): from 2395261 to 2613730. NID: g2634723. >gp:gp|L47648|B

25 ACSERA_11 Bacillus subtilis phosphoglycerate dehydrogenase (
 serA), ypaA, ferredoxin (fer), ypbB, recS, ypbD, ypbE, ypbF,

ypbG, ypbH, glutamate dehydrogenase (ypcA), ypdA, ypdB, ypd
 C, spore cortex lytic enzyme (sleB), ypeB, ypfA, ypfB, cytid

ine monophosphate kinase (cmk), ypfD, ypgA, yphA, yphB, yphC
 , NAD+ dependent glycerol-3-phosphate dehydrogenase (glyc),

30 yphE and yphF genes, complete cds. NID: g1146195. >gp:gp|L47
 648|BACSERA_11 Bacillus subtilis phosphoglycerate dehydrogen

ase (serA), ypaA, ferredoxin (fer), ypbB, recS, ypbD, ypbE,
 ypbF, ypbG, ypbH, glutamate dehydrogenase (ypcA), ypdA, ypdB

35 , ypdC, spore cortex lytic enzyme (sleB), ypeB, ypfA, ypfB,
 cytidine monophosphate kinase (cmk), ypfD, ypgA, yphA, yphB,

yphC, NAD+ dependent glycerol-3-phosphate dehydrogenase (gl
 yc), yphE and yphF genes, complete cds. NID: g1146195.

40 gtgcgcatccagtagcgaatggacgatgggacagttaaaacatttactggataccgtgca
 caacataatgatgcagttggccctactaaagggtggcggttcggttccatccagaagtgtat

gaagaagaagtggaagcattatcaatgtggatgactttaaaatgtggcattgttar:ttta
 ccttatggtgggggttaaaggcggtattgtttgtgatccacgtcaaagtagtatac:cgaa

45 gttgagcgtctatctcgaggggtatgtaagagctatatctcaatttgttgggcctaacaaa
 gatattccagcaccagacgtttttacaaactctcaaattatggcttggatgatggatgag

tatagtgacattagataaaatttaattcgccaggttttataacaggcaaacctattgtatta
 ggtggatctcaaggacgtgatcgctctacagcttttaggtgtagtcattgcaattgaacaa

50 gcgcaaaaacgctcggtggcatggatattaaagatgccaaaattgtgattcaagggttcggc
 aatgcaggtagtttcttagctaaattcttatacagatttaggtgctaaagtagttggata

agtgatgcataatggagccttacatgaccctaattggacttgatatagattatttattagat
 cgacgtgatagcttcggtacagttacaaatctattcgaagatacaatttctaacaaagaa

55 ttattcgaattggattgtgacatccttggtcctgctgcaatttccaatcaaatcactgaa
 gataatgcgcacgatattaaagcaagcattgtagttgaagctgctaattggccctacgacg

cctgaagcaacacgtatttttaacagaaaagagatatattactagtgccgatgtacttgca
 agtgacaggaggtgtgactgtatcttatttcgagtggtgacaaaataatcaaggttattat

60 tggactgaagaagaagtcattgacaaaactacgtgagaagttagtaacagcatttgatacg
 atttatgaattgtcacaaaatagaaaaattgatatgagattagcagcatatatagtaggt

attaaacgtactgctgaagcagcaagatatagaggttgggcataa

Sequence 2162

VRIPVRMDGTVKFTFTGYRAQHNDVAGPTKGGVRFHPEVDEEEVKALSMWMTLKCGLVNL

PYGGGKGGIVCDPRQMSIHEVERLSRGYVRAISQFVGPNKDI PAPDVFTNSQIMAWMMDE
 YSALDKFNSPGFITGKPIVLGGSQGRDRSTALGVVIAIEQAARKRGMDIKDAKIVIQGGF
 NAGSFLAKFLYDLGAKVVGISDAYGALHDPNGLDIDYLLDRRDSFGTVTNLFEDTISNKE
 LFELDCDILVPAAISNQITEDNAHDIKASIVVEAANGPTTPEATRILTERDILLVPDVLA
 5 SAGGVTVSYFEWVQNNQGYWTEEEVNDKLEKLVTAFTDIYELSQNRKIDMRLAAYIVG
 IKRTAEAAARYRGWA*

Sequence 2163

Contig_0718_pos_4955_5407,

10 putative peptide of unknown function

atgtatgatgaattagaggttaataaaagtcgttttaaaaactgtaattttaatgaaggt
 atttttaagaatatagaagcaatttgtaattgtataattacaacgtgcgggtttaataat
 tgtattttcgaagatgttcatttttacaacaaacatttaagattcaacatttgatgaat
 acaccatttgatcaatccgtatttaatagtactttattccaaaatgcaatgttcgatagc
 15 aatctcattcgtagcgtaaaatggactgatatcatttttaaaaacgtttctttcaaaaat
 gtagaaattgaaggaacaacatttaagatgtaaaattcaaaaattgtgagttcaaaaat
 gtaattattactaattcaactatgtcgcgaaaagttaataatgaattacaaaacaagat
 gttacttttagaaaatatagacacttctatttaa

20 Sequence 2164

MYDELEVNKSFRKNCNFNEGIFKNIEAICNCKFTTCGFNNCIFEDVHFYKNQFKDSTFVN
 TPFQDSVFENSTLFQNAMFDSNLIRSVKWTDIIFKNVSFKNVEIEGTTFKDVKFKNCFEKN
 VIITNSTMSQKLMNELQKQDVTLENIDTSI*

25 Sequence 2165

Contig_0718_pos_8056_7316,

is similar to (with p-value 3.0e-34)

>sp:sp|Q06174|EST_BACST CARBOXYLESTERASE PRECURSOR (EC 3.1.1
 .1). >pir:pir|JC1374|JC1374 carboxylesterase (EC 3.1.1.1) -
 30 Bacillus stearothermophilus (strain IFO 12550) >gp:gp|D12681
 |BACPBH7_1 Bacillus stearothermophilus esterase gene. NID: g
 2163:3.

gtgaatgtaaaaatgaaagttaaatcaccacaatcaatctactttaaaggacatcgtaa
 caagctgtattgttattacattcttttacgggaactgtacgtgatgtaaaacatttagca
 35 caacagtgtaatgaagagggatttacttgttacgtgcctagtatccaggccacggttg
 ccacttaaggaatttacccaacacaatattaatgattggtgggaacaagttacagcagca
 tatcaatttttaagaaatgcaggatacagtagaattaatgtgacaggcgtatcattaggg
 ggattatttactttaaggttagctgaacattttgatttagaacgtatagctgtgatgtca
 gccccacataaaaagcgtgaaagcgagattgcgtggcgctcttgaaaggatgggcatoga
 40 atgaatgaaattttgagtttaagcgaagaagagcgtcgtcaccaaatggaaaccatcttg
 tcttatgataaagaaattgaagtgtttcaaggtgtaattgatgaaattatggcttatctt
 gcaaatattacagtaccagtgaatattatgtatggcgaagaagatgacccatttatgct
 caaagtgcgcaatacatttatgataatgtaaatagtcagataaagaactgctcaaat
 gaaaaaagcggatcatcttatgacgtatggcgatcatgcatacagagtagaacaatctatt
 45 attcaatttttcagtaataaa

Sequence 2166

VNVKMKVKSQSIYLGHRQQAVALLLHSFTGTVRDVKHLAQQLNEEGFTCYVPSYPGHGL
 PLKEFTQHNINDWWEQVTAAYQFLRNAGYSRINVTGVSLLGLFTLRLAEHFDLERIAVMS
 50 APHKKRESEIAWRLERYGHRMNEILSLSEEERRHQMETILSYDKEIEVFQGVIDEIMAYL
 ANITVPVNIMYGEEDDPLYAQAQYIYDNVNSQDKELLKFEKSGHLMTYGDHAYRVEQSI
 IQFFEK*

Sequence 2167

55 Contig_0718_pos_6737_5586,

putative peptide of unknown function

atggaatttacggttaggtaaaatgggtcgtacatatacaacgcaaatatataagaaatta
 acgggaaagaaatggcttaattatcggtggaatggtaatttagccgtatttatacta
 tttggtttttatagtgttattggtggttgattatttatatatataggttatgtcatagca

caaatcatgggtttttaaatcaagtagcgtgacaaatattcaatttgaaacaatcattagt
 aatccatgggtgactgttttaggtcaaggcatattttttgataacaatggtaattggt
 atgttaggtgttgaaaaaggtttagaaaaagcttctaaaataatgatgcctctattattt
 atctttt.aattatcggttagcacaaatctttaactttagaaggtgctttagaaggtgta
 5 cgttatatactgcaacctcgagttgaagatatgtctattcaaggtgactatttgcgtta
 ggacaatcggtttttacgctgtccctagggtacaaccggaatgattacttatgcaagctat
 gcacctaaaaatagacgataaaagtcttcagcactttcaattgtcgaatgaatatttta
 atttctgtcttggctggattagctatatttcctgctgcttaaacatttggttaccaaccc
 caagaaggccctggcttattatttaagggtttaccactagtagtttagcgaaatgactttt
 10 ggtacattcttttactttatatttttactattattcttatttgcggcattaacgtcttct
 atatcattattagagttaaatgtatctaattttactaaaaatgataatagtaaaagacaa
 aaatgggcaatcataggtagtatacttatttatcatttagtatccagcaacattatct
 ttagtagtctaagtcatttgcgttttggcgctggtacgatatttgataataggtttt
 attgtatctaataattcttatgccattaggggactaggaacaacattagtggttgccaa
 15 ttactagataaaaaattattaaagaaagctttgggaaagacaaattcaacctattttta
 ccgtggtattatttaattaagttcatcatgcctattgttattatttttagtatttatagtt
 caattatttttaa

Sequence 2168

20 MEFTVKGMRITYTTQIYKKLTGKKWLNIIIGWNGNLAVFILFGFYSVIGGWIIIIYIGYVIA
 QIMVFKSSTLTNIQFETIISNPWLTVLGQGFILITMVIVMLGVEKLEKASKIMMPLLF
 IFLLIIVVAQSLTLEGALEGVRYILQPRVEDMSIQGVLFALGQSFFTLGLTGMITYASY
 APKNMTIKSSALSIVVMNILISVLAGLAIFPALKTFGYQPQEGPGLLFKVLPLVFSEMTF
 GTFFYFIFLLFLFAALTSSISLLELNVSNFTKNDNSKRQKVAIIGSILVFIISIATLS
 25 FSSLSHLRFAGTIFDNMDFIVSNILMPLGALGTTLVVGQLLDDKLLKESFGKDKFNLFL
 PWYYLIKFIPIVILVFIIVQLF*

Sequence 2169

Contig_0718_pos_4565_3657,
 30 is similar to (with p-value 9.0e-76)
 >gp:gp|U93874|BSU93874_1 Bacillus subtilis cysteine synthase
 (yrhA), cystathionine gamma-lyase (yrhB), YrhC (yrhC), YrhD
 (yrhD), formate dehydrogenase chain A (yrhE), YrhF (yrhF),
 formate dehydrogenase (yrhG), YrhH (yrhH), regulatory protei
 35 n (yrhI), cytochrome P450 102 (yrhJ), YrhK (yrhK), hypotheti
 cal protein YrhL (yrhL), putative anti-SigV factor (yrhM), R
 NA polymerase sigma factor SigV (sigV) and YrhO (yrhO) genes
 , complete cds, and YrhP (yrhP) gene, partial cds. NID: g193
 4604. >gp:gp|Z99117|BSUB0014_206 Bacillus subtilis complete
 40 genome (section 14 of 21): from 2599451 to 2812870. NID: g26
 34966.
 atgattgcatacgaatttgataggacaaactccattagttttattagaagcttttagtgac
 gagaatgttaaaatatacgccaaacttgagcaatttaatcctggtggttagcatcaaaagac
 cgtctagggaagtacttaattgaaaaagcaatagatgaaggacgacttaagaaggggat
 45 acaatagtcgaagcgactgctgtaatacaggcattggacttgctattgcttctaactcg
 caaaaagtaaaatgtatcatctttgctccagaaggatttgcagaagaaaaatttcaatt
 atgaaagcattgggtgcagatgttagacgtacccccaaagctgagggaaatgactggcgca
 cagcaagaggcggttggcatacgaacacgatattggatatttatatatgaatcaattcgaa
 actaaagataatcctggggcatatacacaacacttgccaaacaactcacagatgaactt
 50 tcacatatgtatttttgtggcaggtgttgggtccggtggtacgtttacaggagttgca
 caacacttaaaaacgtatgatgtaaaaaattatattgtagaaccagaaggctctgtctta
 aatggtgtgtgcagtcactcctcatgcaactgaagggttgggttctgaaaagtggccatca
 tttttagaaaaagaattagtagatgggtatttttactggttgcgataaagatgctttta
 aatgtttaaacttgtcgcgaataaagaaggattgttagttggttagttcttcgggagcggca
 55 ttacaaggagcggttgaattaaaaaaaagcattcaaaatggtgtgattgttaccatctt
 ccagatggaagcgatcgatacatgtccaaacaaatattcaactataaggagagttttaat
 aatgaataa

Sequence 2170

MIAYDLIGQTPVLVLESFSDENVKIYAKLEQFNPGGSIKDRLGKYLIEKAIDEGRLKEGD
TIVEATAGNTGIGLAIASNRHKVKCIIFAPEGFAEEKISIMKALGADVRRTPKAEGMTGA
QQEALAYATRYGYLYMNQFETKDNPGAYTQTLAKQLTDELSHIDYFVAGVSGGTTFTGVA
QHLKTYDVKNYIVEPEGSVLNNGGVSHPHATEGIGSEKWPSFLEKELVDGIFTVADKDAFN
5 NVKLVANKEGLLVGSSSGAALQGALELKKSIQNGVIVTIFPDGSDRYMSKQIFNYKESFN
NE*

Sequence 2171

Contig_0718_pos_3646_2519,
10 is similar to (with p-value 0.0e+00)
>gp:gp|U93874|BSU93874_2 Bacillus subtilis cysteine synthase
(yrhA), cystathionine gamma-lyase (yrhB), YrhC (yrhC), YrhD
(yrhD), formate dehydrogenase chain A (yrhE), YrhF (yrhF),
formate dehydrogenase (yrhG), YrhH (yrhH), regulatory protei
n YrhI (yrhI), cytochrome P450 102 (yrhJ), YrhK (yrhK), hypothe
15 cal protein YrhL (yrhL), putative anti-SigV factor (yrhM), R
NA polymerase sigma factor SigV (sigV) and YrhO (yrhO) genes
, complete cds, and YrhP (yrhP) gene, partial cds. NID: g193
4604. >gp:gp|Z99117|BSUB0014_205 Bacillus subtilis complete
20 genome (section 14 of 21): from 2599451 to 2812870. NID: g26
34966.

atgatacatgggggacatacgcagacaactataactggagcagtgacaacacctatttat
caaacaagtacttatttacaagatgatattggtgatttaagacaaggtacgaatattca
cgtactgcaaatcctacacgtgcgtctcttgaaagtgttattgctaatttagaacatggt
25 aagcatgggttttgcttttggttcaggaatggcagcaattagtgacgttatcatgttatta
gataaaggagatcacttagttcttaattctgatgtttatggtggcacatatcgtgcatta
actaaagtatttactcgcttttggtatagacgttagattttgttgatacaactaaaattgaa
aacattgaacaatatattaaacctgaaactaaaatgttatatgtagaaacaccttcaaat
ccattattgcgtgtgactgatattaaagcatcagcaaaaattgcaaaaaaatatgatttg
30 atatctgtagtcgataatacatttatgacaccttactacaaaaccccttagactttggt
attgatatcgattgcattcggtacaaaatatattggaggccatagtgatgtttagct
ggtcttgttgctactgctgatgatgatttagcagaacgtctaggctttatttcaaattct
acagggtggtgtacttggacctcaagatagctatttattaatcagagggtattaaaacgcta
ggtctaaagaattggagcaataaaacgaaacgttgaaggattgtgcaaatgttacaaaag
35 caccctaaagttcaacaagtattccatcctagttattaaggaacatatgaactatactatc
catcaaaatcaagcaactgggcatacaggggtagtatcttttgaaagttaaagatacagaa
gcggctaaacaagtgattcacgcaaaaactactttacactggcagagagtttaggggca
gttgaaagtctaatttctgtaccggcacttatgacgcatgctccatcccatcagatgta
agagccaaggaaggtattacggatggtctcattcgtttatctattggtattgaagacaca
40 gaagacttagttaatgatttagaacaagccctaaatactttgagataa

Sequence 2172

MIHGGHTDNYTGAVTTPIIYQTSTYLQDDIGDLRQGYEYSRTANPTRASLESVIANLEHG
KHGFAFGSGMAAISAVIMLLDKGDHLVLNSDVYGGTYRALTKVFTRFGIDVDFVDTTKIE
45 NIEQYIKPETKMLYVETPSNPLLRVTDIKASAKIAKKYDLISVVDNTFMTPIYQNPIDFG
IDIVLHSATKYIGGHSVDVAGLVATADDDLAERLGFISNSTGGVVGQDSYLLIRGIKTL
GLRMEQINRNVEGIVQMLQKHPKVQVVFHPSIKEHMNYTIHQNQATGHTGVVSFEVKDTE
AAKQVIHATNYFTLAESLGAVESLISVPALMTHASIPSDVRAKEGITDGLIRLSIGIEDT
EDLVNDLEQALNTRL*

50

Sequence 2173

Contig_0719_pos_3554_2802,
is similar to (with p-value 1.0e-71)
>sr:sp|P29928|SUMT_BACME UROPORPHYRIN-III C-METHYLTRANSFERAS
55 E (EC 2.1.1.107) (UROGEN III METHYLASE) (SUMT) (UROPORPHYRIN
OGEN III METHYLASE) (UROM). >pir:pir|A42479|A42479 S-adenosy
l-L-methionine uroporphyrinogen III methyltransferase - Baci
llus megaterium >gp:gp|M62881|BACCOBA_1 Bacillus megaterium
S-adenosyl-L-methionine:uroporphyrinogen III methyltransferas

e (COBA) gene, complete cds. NID: g142694.

gtgggtcatatggggaaagtatatatttagttggagctggacctgggtgatccagaattaata
acgttaaaaggttttaaagccattaaagaagccgatgtcatcctttatgaccgacttgta
aataaagaataacttaattatgcttctccttctactaagttccttctattgcggttaaggat
5 cctcacaggcactccttaccgcaggaagaaacaaataaaatgatggtaaccttagccaaa
aaagggcacatagttacacgttttaaaggggtggcgatccatttgtttttggacgtggcgga
gaagaagcagaggaattagcatgtcataatatccactttgaaattatacctggaattcca
gtaacacatcgtgattatagttccttctgtagcatttgaactgcagtgaataaacctgggt
atggataaaggcaataactggcaacatttggccaatggctcctgaaactttatgtatttat
10 atggggggttaagagactcagtgaaatttgtgagttgtaataacaatatggctcgttcgtca
gaaacaccagtagctctcgtgcatatgggaacgtcaaacagcaaatgcagtgactggg
acactcgatcaaatcaagaacgagcacatcatattcagaatccagcaatgattattgta
ggcgaagtgggttaagatgagagaaaaaatttaattggttgtagaacaggcaactgttcaa
aatgaaacgttaacggaaatgtcatcaacttag

15

Sequence 2174

VVHMGKVYLVGAGPGDPELITLKLKAIKEADVILYDRLVNKEILNYASPSTKFFYCGKD
PHRHSLPQEETNKMMVTLAKKGHIVTRLKGGDPFVFGRGEEAEELACHNIHFEIIPGIP
VTHRDYSSSVAFVTVAVNKPMDKGKYWQHLANGPETLCIYMGVKRLSEICELLIQYGRSS
20 ETPVALVHMGTSKQQMTVTGTLDTIQERAHHIQNPAMIIVGEVVKMREKINWFVEQATVQ
NETLTEMSS*

Sequence 2175

Contig_0719_pos_2721_2116,
is similar to (with p-value 4.0e-19)
25 >gp:gp|AJ000974|BSPYREYLO_8 Bacillus subtilis pyrE to yloA g
ene region. NID: g2462954. >gp:gp|Z99112|BSUB0009_34 Bacillu
s subtilis complete genome (section 9 of 21): from 1598421 t
o 1807200. NID: g2633902.

atgcccttaatgattgatttaagtaacaagaaagtcgtcattgtaggtggaggtaaagt
gcaacacgtcgtgctaaaaactttattagcttatacaaaacatattcatgttgtaagtcca
acaattaccgatacattacaaaaatatctagaacgaagcaaatcacttatgaaaagaaa
cacttcgaaccacaagatggtgagaatgctgatgtggtcatcgcggtactaatcaatct
gatgttaacaacgatgtgggggcagctttgtctaagaacgtattatttaatcatgcagga
35 caagcagacctaggtaatgtaacgttccttaatttcttaaaaagagataaattaacaata
agtgtatcaactgatggtgcaagtcctaaatttaggtcaacgaattattaaagatttaaaa
gatacatacaatgaagactattcaatgtatattcagtttttatatgaaagtagacaatat
attaaatcacttaaaattgagccatctgataaacaagcgttactcgagcaaattttgtca
gacaaatatttagatgagaagaagcaacaagatttcatccgatggctaaaatcacaagtc
40 aatga

Sequence 2176

MPLMIDLSNKKVVIVGGGKVATRRAKTLLAYTKHIHVVSPTITDTLQKYLETKQITYEKK
HFEPQDVENADVIAATNQSDVNNDVGAALSKNVLFNHAGQADLGNVTFPNFLKRDKLT
45 SVSTDGASPKLGQRIKDLKDTYNEDYSMYIQFLYESRQYIKSLKIEPSDKQALLEQILS
DKYLDEKKQQDFIRWLKSQVK*

Sequence 2177

Contig_0719_pos_2029_1196,
putative peptide of unknown function
atgggatttggcgcttcatcgtcatcaatattattaacttacggatatagcaccggcagta
gtgtcagcaaccgttcatttttctgaaattgcaacaacagctgcatctgggacatcacat
tgagatttgataatggtcataaaaccaaatggtgaagttagctatacctgggtcaata
agcgccctttatcggtgcaggtgttttgacatttattcatggtgattatattaaaccattc
55 attgctttattcttgtaagtattgggattttatatttggatcaatttctatttaaacgt
gcacatgaacatcatcatcatgtgggaaatttgagtagttttaagtaattccacaaggt
tttgtggcaggatttttagacgcaatcggtgggtggtgggtggggaccggttaatacgcgg
ctcctgctttcaagtaaaaaaattcaaccacgatatgcgattggaacagtcctcagcaagt
gaattttttgttacgtcatctgccgctttaagtttcattatcttttttaggagtcactcaa

attaattggtttgctgtaattgctttaagctctcggtggaatggtagcagcacctatttca
gcgtatttagttaaagtgttaccattaacattcttgcaatttggtcggtggttaatt
atatttacaaatagtaatgcattattaagctattttgtaaaagataacactatttcaa
acagttcgattcattattattcttgcaattattattttgcttggttttcaagtcggtcga
5 aacaagaaattgtctttttcttataagaaaagccgagtaacaaatataattaa

Sequence 2178

MGFGASSSSILLTYGIAPAVVSATVHFSEIATTAASGTSWRFNDNVHKPTMLKLAIPGSI
SAFIGAGVLTFIHGDYIKPFIALFLLSMGFYILYQFLFKRAHEHHHHVGNLSSFKVIPQG
10 FVAGFLDAIGGGGWGPVNTPLLLSSKKIQPRYAIGTVSASEFFVTSSAALSFIIFLGVTQ
INWFAVIALSLGGMVAAPISAYLVKVLPI NILAICVGLLIIFTNSNALLSYFVKDNTISN
TVRFIIILAI IILLVFQVVRNKKLSFSYKKS RVNKYN*

Sequence 2179

15 Contig_0719_pos_0_1172,
is similar to (with p-value 0.0e+00)
>gp:gp|AJ000974|BSPYREYLO_4 Bacillus subtilis pyrE to ylbA g
ene region. NID: g2462954. >gp:gp|Z99112|BSUB0009_30 Bacillu
s subtilis complete genome (section 9 of 21): from 1598421 t
20 o 1807200. NID: g2633902.
atgtctaacaatgaaacaataaccaattatacaattaaacctcatggaggagaactcatc
aatcggtgtgtgaaggaaacgaacgtgaacgtttgattgaggagcattaaattttaa
ccgattactttaaatccttggggaatatcggtatcagagctcataggtattggcggattt
agtcaccttacaggatttatgaacaaggaagactacactaaggttatagaggaaacacat
25 ttaagcaatggcttagtttgagatttctctatcactttacctgtacagaatccgaagca
gataaacttgaaataggtgatgatattgctttatatgggtgaagatgggtcagttatatgga
acgcttaaattagaagaaaagtacacatatgataaagaaaaagaagcgctttggtgtac
ggaactactgaagaagctcatcctggagttaaaaaggtttatgaaaaaggtaatatatat
ttaggtggtcctattaaactattaaatcgctccaaacatgacgcgttttcaaattatcat
30 ctggatccttcagagacgagacaattatttcatgatttaggttgaaaactgtcgtaggt
tttcaaacgagaaatccagtgcatcgagcacatgaatatattcaaaaatcagcactagaa
attggttgatggcttacttttaaatccactagttggtgaaacaaagtcagacgatattcca
gcggatgtacgtatggaaagttatgaagtgatattaaaaaactattatcctgaagataga
gcacgtctagtcattttatcctgctgcaatgcgctatgccggaccacgtgaagcgatactt
35 catgcaactgtccgtaaaaattatggttgacacattttattgtgggaagagatcacgct
ggggtagcgattattatggtacttatgaagcacaagagctgattactcaatttgaagat
gagttaggtattcaaattttaaaatttgaacatgccttttattgcgaagcttggtgaaat
atggcaactgctaaaacatgtccgcatgacgcttctcaacatttacatttaagtgtact
aaagtaagagaaaaactgcgtaatggcgaatcattgccaaactaaattttcaagaccagaa
40 gttgccgaagttctaattaaaggtttgcgagT

Sequence 2180

MSNNETITNYTIKPHGGELINRVVEGNERERLIEEALNFKPITLNPWGISDLELIGIGGF
SPLTGFMNKEDYTKVIEETHLSNGLVWSIPITLPVTESEADKLEIGDDIALYGEDGQLYG
45 TLKLEEKYTYDKEKEARLVYGTTEEHPGVKKVYEKGNILYLGPIKLLNRPKHDAFSNYH
LDPSETRQLFHDGLGWKT VVGFTQTRNPVHRAHEYIQKSALEIVDGLLLNPLVGETKSDDIP
ADVRMESYEVLKNYPEDRARLVIYPAAMRYAGPREAILHATVRKNYGCTHFIVGRDHA
GVGDYYGTYEAQELITQFEDELGIQILKFEHAFYCEACGNMATAKTCPHDASQHLHLSGT
KVREKLRNGESLPTKFSRPEVAEVLKGLRV
50

Sequence 2181

Contig_0720_pos_409_906,
is similar to (with p-value 3.0e-25)
>sp:sp|P35164|RESE_BACSU SENSOR PROTEIN RESE (EC 2.7.3.-). >
55 pir:pir|S45560|S45560 hypothetical protein X18 - Bacillus su
btilis >gp:gp|L09228|BACDIA_27 Bacillus subtilis spoVA to se
ra region. NID: g410114. >gp:gp|Z99116|BSUB0013_23 Bacillus
subtilis complete genome (section 13 of 21): from 2395261 to
2613730. NID: g2634723.

atggatgctgaaggattatcagttgagaaggaattacaacctattcaacaccttcttgat
 aaaatggagtctaaatatcgcatgcaaagtgaagaattaggtttaacaatgacgtttgat
 tctaataatgacgaacaattatggaactatgatatggatagaatggaccaagtgttaact
 aatttaattgataacgcaacaagatatacacaagctgggtgattctataaagattttctatt
 5 gatgaagattcagattttcaatattttaacaataactgatacaggcactggatagcaccg
 gaacatctgaaacaagtatttgaccgtttttataaagtgacgctgctcgaaaaagaggt
 aagcaaggcaccggattaggacttttcatattgtaaaatgattattgaagaacacggggga
 cgtattgatgttgagagcgaattaggcaaggtacttcatattatttagactaccta
 10 tcaaacaaattagttag

Sequence 2182

MDAEGLSVEKELQPIQHLLDKMESKYRMQSEELGLTMTFDSNNDEQLWNYDMRMDQVLT
 NLIDNATRYTQAGDSIKISIDEDSDFNILTITDTGTGIAPEHLKQVDFRYKVDAARKRG
 KQGTGLGLFICKMIIIEHGGRIDVESELGKGTSFIIRLPKSKQIS*

Sequence 2183

Contig_0720_pos_1420_2037,

is similar to (with p-value 7.0e-39)

>sp:P50726|YPAA_BACSU HYPOTHETICAL 20.5 KD PROTEIN IN SER
 20 A-FER INTERGENIC REGION. >gp:gp|Z99116|BSUB0013_17 Bacillus
 subtilis complete genome (section 13 of 21): from 2395261 to
 2613730. NID: g2634723. >gp:gp|L47648|BACSERA_2 Bacillus su
 btillis phosphoglycerate dehydrogenase (serA), ypaA, ferredox
 in (fer), ypbB, recS, ypbD, ypbE, ypbF, ypbG, ypbH, glutamat
 25 e dehydrogenase (ypcA), ypdA, ypdB, ypdC, spore cortex lytic
 enzyme (sleB), ypeB, ypfA, ypfB, cytidine monophosphate kin
 ase (cmk), ypfD, ypgA, yphA, yphB, yphC, NAD+ dependent glyc
 erol-3-phosphate dehydrogenase (glyc), yphE and yphF genes,
 complete cds. NID: g1146195. >gp:gp|L47648|BACSERA_2 Bacillu
 30 s subtilis phosphoglycerate dehydrogenase (serA), ypaA, ferr
 edoxin (fer), ypbB, recS, ypbD, ypbE, ypbF, ypbG, ypbH, glut
 amate dehydrogenase (ypcA), ypdA, ypdB, ypdC, spore cortex l
 ytic enzyme (sleB), ypeB, ypfA, ypfB, cytidine monophosphate
 kinase (cmk), ypfD, ypgA, yphA, yphB, yphC, NAD+ dependent
 35 glycerol-3-phosphate dehydrogenase (glyc), yphE and yphF gen
 es, complete cds. NID: g1146195.

atgggagaagatggaggtttttgtgttcaaatttcctccttttctaagtaagatgaat
 ggaaggagaaaattatatatgcaacaaaacaaacgtttgattacaattagttatgttaagt
 gcggtagcgtttgtgttaactttcatcaagtttccattgccatttataaccaccgtatcta
 40 actctcgatttttagtgatgtaccgacgttatttagcaacattcctcttaagtcctattgct
 gggattatcgttgactcatcaaaaatattttaattttctattcaatataggggaccc
 gttggaccagtagctaacttttttagcaggcgtcagctttttgctatcatcactatggt
 tataaaaaagaaaaataatcgttctttaatttatggattaattacaggtacaatcggt
 atgactattgttttgagcatcttaattattttgtgttacttccattatatggaatgata
 45 ttttaatttaggtgatgtgcttaataacgtaaaaattgttattgtgtctggagtcatacct
 ttttaatttaattaaaggcataatcatttccattatatttgtgctgttatttagaagatta
 agacatatcatcaataaa

Sequence 2184

MGEDGGFLLFKFPFLSKMNGRRKLYMQQNKRLITISMLSAVAFVLTFIKFPLPFIPPYL
 50 TLDFSDVPTLLATFLLSPIAGIIIVALIKNILNLFNIGDPVGPVANFLAGVSFLLSSYYV
 YRKRKNNRSLIYGLITGTIVMTIVLSILNYFVLLPLYGMIFNLGDVLNNVKIVIVSGVIP
 FNLIKGIISIIIFVLLFRRLRHIK*

Sequence 2185

Contig_0720_pos_2823_3761,

putative peptide of unknown function

atgagtcatgcatttaactacaaaaccaataaaaagtatctataatattttaacaggcaag
 aagtcacaccaaacgtttttcgatgcgtcaagccaacaacttttgtcatttatatcatagt

ttacctaacttaagatttcaacttttgagcaatttatactccaaaaggatgattttaa
 aaatcaattcaagtaaaaaatacatccacaatacacttatgatagtctcactcaaacttt
 agttgcatacaattacttatccaaacggtatctcatcacgcaaggagtcaaatacattt
 attccaatcggttcaaaaatcctatatccaacaaagggttaaacaactttatcatcaggtc
 5 attgaatcaaatcaagtatcaaatactatagacgaaatataatttatttttgagaattta
 aataataaataataaccatacatttcttcattatttttacaaggatatgaggaatccatg
 tatactagacaacaaataagttaattgagagtataccacaatcagaattatttgaacga
 gaaatgaatgaactgattgacatatgaatcaattaaaagattcaacgaaatatccaata
 ctttctcaagctatcattctttcaccattactaacaatacatacttaagctatcaaaag
 10 ttaaaatctggtctcaattttaaagaaattgctcaattacaaaatgtaaaacttaacaca
 attgaagatcatattctagaatgtatattaaagggttatttgatagactatacattattt
 ataaataaaaaagatattctcgaatttataaactactatcaaaaacatcgcggtgaacga
 ttaaaattttataaagaacattttactgattggacttactttcaaattaagttagtata
 gtaggaatagaaagggtgatttaattgctgaaagataa

15

Sequence 2186

MSHAFNYKTNKSIYNILTGKKSHQTFDASSQQLLSLYHSLPNLKYSTFEQFILQKDDFK
 KSIQVKIHPQYTYDSLTTQTFSCIQLLIQTLSHTRKESNTFPIVQNTYIQQRVKQLYHQV
 IESNQVSNITIDEIYLLFENLNNKYNHTFLHYLQGYEESMYTRQQISLIESIPQSELFER
 20 EMNELIDILNQLKDSTKYPILSQAILSPLLTNTYLSYQKLKSGNLNKEIAQLQNVKLNT
 IEDHILEMYIKGYLIDYTLFINKKDILEFINYYQKHRGERLKFYKEHFTDWTYFQJ KLVI
 VGIERGDLIAER*

Sequence 2187

25 Contig_0720_pos_5757_5161,

putative peptide of unknown function

atgtcaatggataaaagctacaaaatatgcacttcaattaagggtgattgctcaagaaagc
 tatectagtgatgacacattcctttctagtcgaaaaattacgcttacttattcaacaa
 ttaaaacaatctattcatcatctcaacaattagatgatgccatgattcaattagcacia
 30 caactcgattattttgaaaatattcattcgatacctggtattggttaagctaagcacagct
 atgattattggggagattggtgatattaagcgatttaaatcaataaaacaactcaatgct
 tttggtggcattgatatcaaacgatatcaatcaggtcatacacactgtagagataccatc
 aacaagcgtggaataaaaaagcgagaaaaactttattttgggtgattatgaatataata
 agaggcgagcatcattatgacaatcatgctcgattattactacaaaactaagaaagcag
 35 cctaatagagaaacctcataagactgccatcattgcttgataaatcgattattaaaaaca
 attcattatcttgtaataatcataaattgtacgattatcaaatgtcaccacattag

Sequence 2188

MSMDKATKYALQLRVIAQESYPSVDRHSFLVEKLRLLIQQLKQSIHHLKQLDDAMIQLAQ
 40 QLDFYFENHSIPGIGKLSTAMIIGEIGDIKRFKSNKQLNAFVGIDIKRYQSGHHCRTI
 NKRGNKKARKLLEFWVIMNIIIRGOHHYDNHVVDYKYKLRKPNEKPKHTAIACINRLKLT
 IHYLVNMHKLVDYQMSPH*

Sequence 2189

45 Contig_0721_pos_858_1772,

is similar to (with p-value 0.0e+00)

>sp:sp|P50307|METHK_STAAU S-ADENOSYLMETHIONINE SYNTHETASE (EC
 2.5.1.6) (METHIONINE ADENOSYLTRANSFERASE) (ADOMET SYNTHETAS
 E). >gp:gp|U36379|SAU36379_1 Staphylococcus aureus S-adenosy
 50 lmethionine synthetase gene, complete cds. NID: g1020316.
 atgggttaaaaataaacagtcctgatattgcacagggtgtagacaaagctcttgagtat
 cgaaatgatatttctgaagaagaaattgaagcaacagggtgcagggtgaccaaggattaatg
 tttggatagcaactgatgaaactgatacgtatatgcctctacctatattcctgtcacat
 caacttgctaaacgattggctgatgtacgaaaagatgaaatttttagattatcttcgtcca
 55 gatggaaaaggtacagggtgactgttgaaatggttaagatgacaaaacttagacgtattgat
 accattgtagtttctacacaacatgctgaagatgtagagtttagcacaaattgaaaaggac
 attaaaacgcatgttattttacccaactgtagataaagctttatttagatgatgaaactaaa
 ttttacattaacccaactggacgtttcgttattggaggacctcaaggagatgctgggtta
 actggacgtaaaattatagttgatagctatgggtgttatgcccgctcatgggtggagg:tgt

tttagtggttaaagatcctactaaagtagatcggttcagcagcttatgcagcaagatatgta
gctaaaaatattgttgcagctgggttagctaaacaatgtgaagtacaacttgcataatgca
attgggtgtagcagaacccggtttccatttcaattaatacgtttgatactggaaagggttca
gaagcacggttagttgaagctgtaagaaagcattttgatttaagaccagcaggtatcatt
5 aaaaatgtagacttaaaacaaccgatataatagacaaacagcagcgtatggcattttgga
cgtagacagcgtattgttaccatgggaaaaattagataaaagtcaatgttttaaaagatgct
gttgaaattcaatga

Sequence 2190

10 MVKNKQSPDIAQGVDKALEYRNDISEEEIEATGAGDQGLMFGYATDETDTYMLPIFLSH
QLAKRLADVRKDEILDYLRPDGKVQVTVEYKDDKPRRIDTIVVSTQHAEDVELAQIEKD
IKTHVIVPTVDKALLDDETKFYINPTGRFVIGGPQGDAGLTGRKIIVDTYGGYARHGGGC
FSGKDPTKVDRSAAYARYVAKNIVAAGLAKQCEVQLAYAIGVAEPVSIINTFDTGKVS
EARLVEAVRKHFRLRPAGIIKMLDLKQPIYRQTAAYGHFGRDVLPLWEKLDKVNVLKDA
15 VEIQ*

Sequence 2191

Contig_0721_pos_1923_2783,

putative peptide of unknown function

20 gtggctgtttttattctttatactcttcttagtagcgaaccatagtaagaagaaagtgaag
aatcaaacagaagcacattataaagaaaaagaacaacatctaaaagaatctcatgaagaa
gcttttagaaaaagagagaggttgagaataaaaaagttgttacaacaaaaagaagatttt
gacgtgacagtttagtaacaaaaatcgtagaattgatgcgttgaaactattctcaaaaaat
catagtgaatatgttacagatatgcgatttaattggtattcgtgagagactagtaaatgaa
25 aagagaatacgcctgaagatatgcataattatggcaaatattttcttgcctagtaattgag
ttaactaatattgaacgtgtgagtcattctgtacttacacgaactggcttatataaatt
gattctcaattattaaaaggccatgtttataatggtatttagtggtgcgcaatttaagaa
ttacctacgatgtcacagtttttgacacgctcgacttagattcatcacagccacagaca
ttggtcttagatcaaaatgaagatcaacattcattatcttttgtaattattcagataag
30 attaaacatattgaaaaattagcaggagattacaaaaatgaattgaatacacaatatatcg
cctacatcaatactgtatttttaattcctaaaaaggataatgatgttacaattttctcatat
acgcagtcacataatgttaaagtttttagttggtcctgaacaattagatgaattcttcaac
aagtttgttttccatggacgcatacagtaaatgtggatgatttacaagatatcatggat
aaaatcgagtcattcaattaa

Sequence 2192

VAVLFFILFLVANHSSKKVKVKNQTEAHYKEKEQHLKESHEEALKEKRVENKVVTKQKEDF
DVTVSNKNREIDALKLFSKNHSEYVDMRLIGIRERLVNEKRIRPEDMHIMANIFLPSNE
LTVNIQVSHVLVTRTGLYIIDSQLLKGVYNGISGAQFKELPTMSQVFDLTLDDSSQPQT
40 LVLQDNEDQHSLSFVNYSDKIKHIEKLADLQNELNKTPTPSILYFNPKKDNDVTISHY
TQSSNVKVLVGPEQLDEFFNKFVHFGRIQYNVDDLQDIMDKIESFN*

Sequence 2193

Contig_0721_pos_3101_3799,

is similar to (with p-value 2.0e-56)

>gp:gp|AL034443|SC4B5_1 Streptomyces coelicolor cosmid 4B5.

NID: g4007668.

atgatttatggttaacgaagaaacggtaggtcaaggatttaagaaggattagaatcgact
ggtttaagtcgtgaagattttatttataacttctaaattatggctaactgattttggacgt
50 caaaatgtggaagatgcctatcgacaatctgttgcaaaattaggactggattatttagat
ttatctatcgatgcattggccaggtacaaatgaagcggtaatgattgatacttggacagggt
atggaagacttgtataaacaaaaatcaggtgaaaaatattggtgtgagcaattttatctcca
gaacattttgaagcattgtctgcccagtttctattaaaccggtcatttaaccaagtagaa
tttcatccttatttaacacaaaatgaattacgaaagtatttagaagctcaaaatattatc
55 atggaatcatggtctccattaatgaattcacaaattctccatgatgaagtcataaatgaa
gtagctaataagtaggaaaaactccagcccaagttgtaataagatggaatattcagcac
gatgttgtgttatacctaaatccgtaacaccacatcgcatagaagaaaatttagacgta
tggaattttgaaattaagcgacaaccaaatggaaagaatcgatcaattaaatcaagataaa
agaattggacctaaccctcttgaatttaacggtaagtaa

Sequence 2194

MIYGNEETVGQGIKEGLESTGLSREDLFITSKLWLTDGFRQNVEDAYRQSVAKLGI DYLD
 LYLHWPPTNEAVMIDTWGRMEDLYKQNVKNIGVSNFTPEHFEALLAQVSIKPV;NQVE
 5 FHPYLTQNELRKYLEAQNIIMESWSPLMNSQILHDEVINEVANEVGKTPAQVVIRWNIQH
 DVVVIPKSVTPHRIEENLDVWNFELSDNQMERIDQLNQDKRIGPNPLEFNGK*

Sequence 2195

Contig_0721_pos_5489_4476,
 10 putative peptide of unknown function
 atggaacgattttgttggtgtaaatcaaattaactatatattcaaatgaatccgtagaagcc
 aaatftaaacgagcgctctaagatcatggaaaactgatcaggcagatgctcataagctt
 gcttggttaggaccgacgcttaacaaacagacagcttacctatacatgagttaatatc
 tttgaattaagagaacgctccggttttcatctagaaatcgagaatgaacaaaatcgactt
 15 aaatftcagatccttgaattactccatcaaacattccctggttttagaaagattgttagt
 agtcgatattcaatcattgcactcaacatcgagaaatctttactcatccagacatggtt
 cttgatatcgacaaggatgtactgattacacatatattcaattctacagataagggaatg
 tcaatggataaagctacaaaatatgcacttcaattaagagtgattactcaagaaagctat
 cctaattgctgatagacattcctttctagtcgaaaaattacgcttacttattcaacaatta
 20 aaacaattctattcatctcaaaacatttagatgatgccatgattcaattagcacaacaa
 ctcgattattttgaaaatattcattcgatacctggtatttggtaaactaagcacagctatg
 attattggggagattggtgatattaagcgatttaaatcaaataaacaactcaatgctttt
 gttggcattgatatacaacgatatacaatcaggtcatcacactgtagagataccatcaac
 aagcgtggttaataaaaaagcgagaaaacttttattttgggtgattatgaatataataaga
 25 gggcagcatcattatgacaatcatgtcgtcgattattactacaaactaagaaagcagcct
 aatgagaaacctcataagactgccatcattgcttgataaatcgattattaaaaacaatt
 cattatcttgtaatgaatcataaattgtacgattatcaaatgtcaccacattag

Sequence 2196

30 MERFCCVNQINYIQMNPLEAKFKTSALRSWKTDQADAHKLACLGPTLKQTDSLPIHELIF
 FELRERVRFHLEIENEQNRLKFQILELLHQTFPGLERLFSSRYSIIALNIAEIFTHPDMV
 LDIDKDVLIHIFNSTDKGMSMDKATKYALQLRVITQESYPNVDRHSFLVEKLRLLIQQL
 KQSIHHLKQLDDAMIQLAQQLDYFENIHSIPGIGKLSTAMIIGEIGDIKRFSKNQLNAF
 VGIDIKRYQSGHTHCRDTINKRGNKARKLLFWVIMNIIRGQHHYDNHVVDYKYKLRKQP
 35 NEKPHKTAIIACINRLKTIHYLVMNHKLYDYQMSPH*

Sequence 2197

Contig_0724_pos_3443_3748,
 putative peptide of unknown function
 40 atgctcctttctgtataactcctataaaaggaggtgaaaatatgaaaagttttattatt
 gcgtatgattttaataacaaaaggattatccaaaattaatagagcgatttgaggattat
 cctaattgttgctaaaatcaataaatcagtttggtttattaattcaactaatgatgctaaa
 actatttgaaacgaattaaaaatgtttattgatagcgatgatagtttggttcggttgtaag
 ctgactggtgaagccgcatggtctaattgtaatttgcagttcacacattttaaagattat
 45 ctttag

Sequence 2198

MLLSAILLIKGGENMKSFIIAYDLNNQKDYPKLIERIEDYPNVAKINKSVWFINSTNDAK
 TIRNELKMFIDSDSLFVGKLTGEAAWSNVICSSQHLKDYL*

50

Sequence 2199

Contig_0724_pos_4321_4635,
 putative peptide of unknown function
 atgtgcttttcaaaaagaatgaaacaatcaagagaaaaacaaggatgacttttagctgaa
 55 ctaggaagaaaaatcggtaaaactgaagctactgtacaacggttatgaaagcggggaatatt
 aaaaatcttaaaaatgatactattgaaagtatagctactgcattaaatgttaaccctgct
 ttcttgatgggttgatagaagaagttgaggaacaaccacaacatcgtcagcgcacatctt
 gatggtgatttaactgacgaagaatggcaagaattcttgattacgctgaatacataaga
 agtaaaagaaaataa

Sequence 2200

MCFSKRMKQSREKQGMFLAELGRKIGKTEATVQRYESGNIKNLKNDTIESIATALNVNPA
FLMGWIEEVEEQPHRAAHLGDGLTDEEWQEILDYAEYIRSKRK*

5

Sequence 2201

Contig_0724_pos_4639_0,

putative peptide of unknown function

gtgttttatgtggggaaatatgaagatatgttaattgaacatgactatattgaagtcatt
10 gaatgtgataaacttacctaaaagggttatctggtttgtggcttgagatatgattttaatt
aatcgtaacttgcctattacttccaaacttgaaacacttgacagaggaaactcgctcataac
gaacttacatatggaaatatagttgatcaaagtagttttaatcatagaaaaattgaaggt
tatgcacgtagggttagcctatgaaaagttaatccctcttaagatatattgtaaaagcattt
ttgcaaggcattcatgacttgtatgaacttgctaatttttttga

15

Sequence 2202

VFYVGKYEDMLIEHDYIEVIECDNLPKRSLGLWLGDMILINRNPITSKLETLAEELAHN
ELTYGNIVDQSSFNHRKFEGYARRLAYEKLIPKDIVKAFLQGIHDLYELANFFX

20 Sequence 2203

Contig_0724_pos_3437_2028,

is similar to (with p-value 0.0e+00)

>gp:gp|M57689|BACSP00K_5 Bacillus subtilis spo0K operon. NID
: gl43602.

25 atgcaagatttacaagtatatttaattttgaagatttaccagtaagaaaaatagaagtagat
ggagaaccatatttttaggtaaagacgtggcagaaatattaggttacacaagatctgat
aatgcaattagaaatcatgttgatgatgaagataagctgacgcaccaagttagtgcatca
ggtcaaaaacgaaacatggtaatcatcaacgaatctggtttatacagcttaattctttgac
gctgctaaacaaagtaaaaacgaaagatttagaaagaaagctaaacgttttaaacgtccc
30 aaaaaatgcgaaaaataaaggaggtaaaaagtaatgacagaaacgggtattagaagtaaatgat
ttgcacgtttcctttgatattgctgcaggagaagtgcagctgtcagaggcgtggatttt
catttaaaataaaggggaaacggttagccattgttggaagaatctggatctggaaaatctgta
acaactaaggcaattacaaaactttttcaaaaggatacaggaagaataaaaaaggagaa
atrttatttttaggtgaggacttagctcagaaaagtgaagaagaactgatacagctaaga
35 ggtcgagatatttcaatgatatttcaggatcctatgacttctttaaatccaacaatgcaa
atcggaagcaagtcattggaacctttgattaaacataagaaattaagtaaaagcaaaggcc
aagcaagagcattggaaattttgaatttagttgggttacctcgtgctgaaaaacgattt
aaagcttatccacatcaattttcaggaggacaacgtcagagaatagttattgcaatagca
ttggcatgtgacctcaaaatatttaattgctgatgagcctacaactgctttagatgtgaca
40 atgcaggctcaaattttagatcttatgaaagaactacaaaataagattgaaacttcaatt
atctttattacgcatgatttaggcgttgtagcaaatattgcggacaaagtagccgtaatg
tatgggggacagatggttgaaacaggggatgtgaatgaaatattttatgatcctaaacat
ccctatacctggggattgctttctcaatgcctgatttaacaaccagtaatgacacggac
ttaattgcaattccagggtacaccaccagatttacttcatccaccaattgggtgatgctttc
45 gcacgtagaagtcgatatgcttttagatattgatttttaagaagaaccaccttggttcaaa
atttcacccacacattttgttaaatcttggttattagatgcaagagctccaaaagttacg
ccaccttcaatggttcaaaaacgattaagaacaatgccaaagtaattatgaacaaccacat
agagtagagagggtggcttttaatgagtaa

50 Sequence 2204

MQDLQVFNFDLPVRKIEVDGEPYFLGKDVAEILGYTRSDNAIRNHVDDDKLTHQVSAS
GQKRNMVVIINESGLYSLIFDAAKQSKNESIRKKAKREKRPKNKIKEVKVMTETVLEVND
LHVSFDIAAGEVQAVRGVDFHLNKGETLAIVGESGSGKSVTTKAITKLFQKDTGRIKKGE
ILILGEDIAQKSEKELIQLRGRDISMIFQDPMTSLNPTMQIGKQVMEPLIKHKKLSKAKA
55 KQRALEILNLVGLPRAEKRFKAYPHQFSGGQRQRIVIAIALACEPKILIADEPTTALDVT
MQAQILDLMKELQNKIETSIIFITHDLGVVANIADKVAVMYGGQMVETGDVNEIFYDPKH
PYTWGLLSMPDLTTSNDTDLIAIPGTPPDLHPPIGDAFARRSRALDIDFKEEPPWFK
ISPTEHVKSWLLDARAPKVTTPSPMVQKRLRTMPSNYEQPHRVERVAFNE*

Sequence 2205

Contig_0724_pos_1756_1097,

is similar to (with p-value 3.0e-81)

>pir:pir|E38447|E38447 sporulation initiation protein spo0KE

5 - Bacillus subtilis >gp:gp|M57689|BACSP00K_6 Bacillus subtilis spo0K operon. NID: g143602.

atgatarattcaagatccttatgcatctctaaatcctcggttaaaggtaatggatatcgta
gctgaaggaatagatatccacaaacttgctagtagtcagcgtgatcgaaagaaacgtgta
tacgaccttttagaaacagttggttttaggtaaagaacacgcgaatcgttaccacatgag
10 ttttcaggcggacaaagacaacgtattggatcgacgtgcattagctgtagagccagaa
tttattattgcagatgaaccgatatcagcattagatgtatcgattcaagctcaagtcgtt
aatctttttattaaagctacaacgtgaacgtgatattactttattggtttattgctcatgat
ttatcaatggtgaaatattttccgatagaattgcagtgatgcacttcggtaaaattgta
gaaattggaccggctgatgatatttataattatccattacatgattataactaagtcatta
15 ttaagtgccattccacagcctgatcctgatgttgagagaaatcgtaacgtgttttatat
catgaagatgcaacgcttaataagaacgtcaattaaatgaaattagaccacaacattat
gtattttctactcaaaacgaagcagttaaattgaaacaaaagtatggtttgtctgtttaa

20 Sequence 2206

MIFQDPYASLNPRKLVMDIVAEGIDIHKLASSQRDRKKRVYDLLETVGLGKEHANRYPHE
FSGGQRQRIGIARALAVEPEFIIADEPISALDVSIIQAQVNNLLKLQRERDITLLFIAHD
LSMVKYISDRIAMVHFGKIVEIGPADDIYNYPLHDYTKSLLSAIPQDPDVERNRRQVLY
HEDATLNEERQLNEIRPQHYVFSTQNEAVKLKQKYGLSV*

25 Sequence 2207

Contig_0724_pos_1070_234,

is similar to (with p-value 2.0e-51)

>pir:pir|A53310|A53310 pheromone cAD1 binding protein precursor
30 - Enterococcus faecalis plasmid pAD1 >gp:gp|L19532|AD1TR
AC_2 Plasmid pAD1 (from Enterococcus faecalis strain: DS16)
hemolysin bacteriocin (traC) gene, complete cds, traA and traB
genes, 3' end. NID: g388267.

atgaaagggttttaaagtccttaattattttattaagtgatgcataattttatctgcttgt
35 agtaataagcagagttttatattcagaccaggggcaagtttttaggaaggtaatcacacaa
gatatgactacactagatacagctttaattacagatgctgtttctggtgatatagcagct
caagcttttgaaggattatatactttaataaagaagacaaagctgaaccagctattgct
aaatcttttccaaagaaaaagtaattggtggcaaaacacttacgattaatttaagaaaaaat
gcaaaatggtccaatggagattcggtaactgcataatgacttcgtatatgcgtggagaaag
40 gatgtaattcaatgaacggtctctgagtttgcatataatgagcgatataaaaaatgca
gatgaagttaatgcaggtaaaaaatcagtcaggatttgggtatcaaggctataggtaaa
tataaattacaagtagattttagaagacctgtaccttatattaatgaactattagcactt
aatacatttaactcctcaaaatgagaaagttgctaaaaagtttggagaacaatatgtaca
actgctgaaaaagcagtgatcaatggaccatttgaagtaacaaattggaaagtggagat
45 aaaattcaattagttaaaaatgaacaatatgggataagaagaatgtaaaattagataaa
gtgaactataaagtattaaaaagatcaacaagcaggtgcacgttatatgatactggctcg
gtcgtatgatactatgttaagtatactgcacaagcatctagccatcagaagggttag

Sequence 2208

50 MKGFKVLIILLVSVCIILSACSNKQSLYSDQGQVFRKVITQDMTTLDTALITDAVSGDIAA
QAFEGLYTLNKEDKAEPAlAKSFPKKSNGGKTLTINLRKNAKWSNGDSVTAYDFVYAWRK
VVNPKTASEFAYIMSDIKNADEVNAGKKSVDLGIIKIGKYKLQVDLERPVPIINELLAL
NTFNPQNEKVAKKFGEQYGTAEKAVYNGPFVETNWKVEDKIQLVKNEQYWDKKNVKLDK
VNYKVLKDQQAGASLYDTGSVDDTMLSILHKHLVHQKV*

55

Sequence 2209

Contig_0727_pos_3292_0,

putative peptide of unknown function

atgtcttttcttaggaaacacaccgaaattatatttagttatatcatcggtatcgtttca

ctttttacaggtctcattatTTTTtattaacttacctttaatcaaacaattttaaggtgac
 aaaaggttgatacgcagtgatgcataacgtagtggaattcctaataatgcctttttgcccag
 attataaaagtgatgagtaaatttattggtggctttccaattacaagtgccatagtaatc
 atcgtatttggattcttagtgatgctgttaggtcacactttatttagaactattaatac
 5 gattatgacatttcaattttcttttttagttattggcattatgtactttatcattacatta
 ttgctaatagcacaagtgtatggcctttttgctatcgtctttattattccttttacagtt
 catattggttacatagttataaaagatgagttgaaccaagacaatcgaaagaaccattat
 atgtggattattgtaacttatggaatgagttatcttattacccaaatttcgctatatgga
 10 cgtattgacgcaaatgaaattgaatcaattgatattttaagtgtaaatacattcttcatt
 attatgtggttatttaggtcagatggctata

Sequence 2210

MSFLRKHTEIIFSYYIGIVSLFTGLIIFINLPLIKQFKGDKKVDTHVHNVWEFLNAFFAE
 IIKVMSKFIGGFITSIAIVIIIVFGILVMLLGHFLFRTIKYDYDISIFFLVIGIMYFIITL
 15 LLMTQVYGFFAIVFIIPFTVHIGYIVYKDELNQDNRNHYMWIIIVTYGMSYLITQISLYG
 RIDANEIESIDILSVNTFFIIMWLLGQMAI

Sequence 2211

Contig_0727_pos_0_369,
 putative peptide of unknown function
 atgtttccaccatattttatcacgaacagggagtgtaaagagcatgacgcgacagagaatc
 gccattgatattggatgaagtgtctgtgatacattgggtgtgtgttaaagcgggtcaat
 gaacgagcggatttaaataatcaaaatggaatcattaaacggtaaaaaattaaaacatatg
 ataccgagcatgaggggttagtcatggatattttaaaagaacctggattctttagaat
 25 ttagatgtaatgccgcagtgctcaagaagttgtaaaacaactcaatgagcattacgacata
 tacatagccacagcagcgatggatgttccaacctctttcatgacaaatatgaatggtta
 ttCAAATGA

Sequence 2212

MEPPYFITNRECKSMTRQRIADMDDEVLADTLGAVVKAVNERADLNKMESLNGKKLKH
 30 IPEHEGLVMDILKEPGFFRNLDVMPHAQEVVKQLNEHYDIYIATAAMDVPTSFDHKEYEWL
 FK*

Sequence 2213

Contig_0728_pos_6214_6576,
 is similar to (with p-value 4.0e-29)
 >gp:gp|AJ000339|LDGAPPGK_3 Lactobacillus delbrueckii ygaP, g
 ap, pgk, tpi, and ycsE genes. NID: g2624189.
 atgactttctattgttttattccacatcaaagtatttaaaaacttatttatcattaata
 40 gcttttataccaggtaattctttaccttctaagtattctaagtgatgctcctccaccagta
 gagatgtgtgtaaaagtcattctcgaaacctaatgaaattgctgctgcggcagagtcacca
 ccaccaataatagtagtagcgtcttccaatttagcaatagactcacatacacgattgta
 cctttagcaaaattactaaattcgaatacacccataggtccattccatactacagtatgt
 gcaccttgtaattctttattaaataattctactgttttaggtccaatatccattgcttct
 45 tga

Sequence 2214

MTFYCFIPTSKYLKTYLSLIAFIPGNSLPSKYSNDAPPPVEMCVKSSSKPNEIAAAESP
 PPIIVVASSNLAIDSHTPIVPLAKLLNSNTPIGPFHTTVCAPCNSLLNNSTVLGPISIAS
 50 *

Sequence 2215

Contig_0728_pos_6923_6258,
 is similar to (with p-value 2.0e-69)
 >gp:gp|AJ000339|LDGAPPGK_3 Lactobacillus delbrueckii ygaP, g
 55 ap, pgk, tpi, and ycsE genes. NID: g2624189.
 atggttaaaagaaattaaatttattggtggcgtagtgaatgatccacaaaaaccagtagtt
 gctatttttaggtggcgctaaagtttcagataaaattaatgttatcaaaaatttagttaat
 atcgcagataaaatcttaatcggtggcggtatggcttatacatttattaaagcgcaaggt

aaggaaataggtctttcattattggaagaagacaaaattgattttgctaaagacttgta
 gagaataatggcgatcaaatagtagtattacctgtagattgtaaaatcgctaaagaattttct
 aatgatgcaaaaatcactgaagtattctaatgaaatcccttcagatcaagaagcaatg
 gatattggacctaataacagtagaattatttaataaagaattacaagggtgcacatactgta
 gtatggaatggacctatgggtgtattcgaatttagtaattttgctaaagggtacaatcggt
 gtatgtgagtcatttgctaaattggaagacgctactactattattgggtgggtgactct
 gccgcagcagcaatttcattaggtttcgaagatgactttacacacatctctactgggtgga
 ggagcatcattagaatacttagaaggtaaagaattacctgggtataaaagctattaatgat
 aatga

Sequence 2216

MEKEIKFIGGVNDPQKPVVAILGGAKVSDKINVIKNLVNIADKILIGGGMAYTFIKAQG
 KEIGLSLLEEDKIDFAKDLLENNGDQIVLPVDCKIAKEFSNDAKITEVSINEIPSDQEAM
 DIGPKTVELFNKELQGAHTVVWNGPMGVFEFSNFAKGTIGVCESIAKLEDATTIIGGDS
 AAAAISLGFEDDFTHISTGGGASLEYLEGKELPGIKAINDK*

Sequence 2217

Contig_0728_pos_6127_5366,

is similar to (with p-value 2.0e-90)

>sp:sp|P35144|TPIS_BACME TRIOSEPHOSPHATE ISOMERASE (EC 5.3.1
 .1) (TIM). >pir:pir|JQ1955|JQ1955 triose-phosphate isomerase
 (EC 5.3.1.1) - Bacillus megaterium >gp:gp|M87647|BACPGKTIMG
 _4 Bacillus megaterium glyceraldehyde-3-phosphate dehydrogen
 ase (gap), phosphoglycerate kinase (pgk), and triose phosphat
 e isomerase (tpi) genes, complete cds. NID: g143315. >gp:gp
 |M87648|BACTPIPGK_2 Bacillus megaterium triose phosphate iso
 merase (tpi) gene, complete cds, and phosphoglycerate kinase
 (pgk) gene, 3' end. NID: g143759.

atgaagaaacccaattatagccggaaactggaaaatgaataaaacagttcaagaagctaaa
 gactttgtaaacgaattaccaacattacctgatcctaaagaagtagaatcagttattgt
 gcaccaacaatccaattagacgcttttagtaacagctgttaaagatggtaaagcaaaagg
 ttaaaaattggagcacaaaacgcttactttgaagaaagcgggtgcttatactggagaaact
 tcaccagtagcattatctgaattaggtgttaaatatgtagtgattgggtcactcagagcgt
 cgtgactatttccacgaaactgacgaagaagtaaacaaaaaagcgcatgctatcttcaat
 cacggtagtacacattattttgtgttaggtgaatctgatgaagaacgtgaagctggtaa
 gcaataaaaatcgtaggtaatacaagtgaagaaagctgtcgaagggttatcagatgatcaa
 cttaaagaagttgttattgcatatgaaccaatttgggctatcggtactggtaagtcatt
 acatctgaagatgcaaatgaaatgtgtgctcacgtacgtcaaacattagctgacttatct
 agtcaagaggttgctgacgctacacgtattcaatatggtggtagtggttaaacctaataac
 attaaagaatatatggctcaatcagatatcgatggcgctcttgtaggtggcgcatcatta
 aaagttgaagatttcgtacaattgttagaagggtgcaaaataa

Sequence 2218

MRTPIIAGNWMKNKTVQEAKDFVNELPTLPDPKEVESVICAPTIQLDALVTAVKDGKAKG
 LKIGAQNAYFEESGAYTGETSPVALSELGVKYVVIGHSERRDYFHETDEEVNKKAHAFN
 HGMPPIICVGESDEEREAGKANKIVGNQVKKAVEGLSDDQLKEVVIAYEPIWAI GTGKSS
 TSEDANEMCAHVRQTLADLSSQEVADATRIQYGGSVKPNNIKEYMAQSDIDGALVGGASL
 KVDFVQILEGAK*

Sequence 2219

Contig_0728_pos_5363_3846,

is similar to (with p-value 0.0e+00)

>sp:sp|P39773|PMGI_BACSU_2,3-BISPHOSPHOGLYCERATE-INDEPENDENT
 PHOSPHOGLYCERATE MUTASE (EC 5.4.2.1) (PHOSPHOGLYCEROMUTASE)
 (BPG-INDEPENDENT PGAM) (VEGETATIVE PROTEIN 107) (VEG107).

atggcaaaacaaccaactgccttaatcatcttagatgggtttcgcaaatcggtgaaagtga
 catggcaatgcagttgaagcaagcacataaacctaattttgatcgatattatgaaaaatat
 cctacaacacaaatagaagctagtggttagatgtaggtcttctgaagggtcaaatgggt
 aactctgaagtaggacatatgaatattggtgcaggacgcatcgatatcaaagtttaact

5 cgtatttaataaatcgattgaagacggagaattctttgataaactgtattaaata:cgct
 gttaaacatgttaaagacaatggctctgcgcttcattgtattcggattgctttctgatgg
 ggtgtacacagtcattataagcatctatttgctattttagaattagctaaaaagcaagga
 atagataaagtatatgtccacgcatttttagatggctcgtgatgttgatcaaaaatctgct
 10 ttgaaatatatagaggaaactgaagataaaatttaagaattaggtgtaggccaattcgct
 tctgttttcaggacgttattatgctatggaccgtgacaagcgttgggatcgtgaggaacgt
 gcctataatgctattcgttaactttgaaggctcctacatttacttcagctaaagcaggcgtt
 gaagctaattataaaaaatgatgtgactgatgaattcgtcgaaccgtttatagttgaaggc
 caaacgatgggtgtgaacgatggagacgcagtaattctttataatttccgtccagataga
 15 gcagctcaactttcagaaatctttactaataaagcgtttgatggatttaaagttgaacaa
 gtggacaacttatctacgctacattcacgaaatataatgacaatgtagatgctgaaatt
 gtattgaaaaagttgactttaataatacaatcggtgaagttgctcaagataatggcttg
 aaacaattacgtatcgctgaaactgaaaagtatccacatgtaacatactttatgagtggt
 ggacgaaatgaagagtttgaaggagaacgtcgtagactcatcgattctccaaaagtagcg
 acttatgatttaaaacctgagatgagtgcatatgaagttaaagatgcattattagaagag
 20 ttagacaaaggtgacttagatttaattctactgaactttgctaaccagatatggttgga
 catagtggatgcttgaaccaacaattaaagcaatcgaagcagtagatgagtgcttgggt
 gaagtcggtgacaaaattattgatattgggtggctcatgccatcatcactgcagaccacggt
 aactcagatcaagatttaactgatgacgaccaacctatgacgacacacacaactaatcct
 25 gttccagttattgttaactaaagaaggtgttacattaaagagaaactggacgttttaggcgat
 ttagcgccgacattattagattttatataatgttaacaacacatctgaaatgacaggtgaa
 tcaactgattaaacattag

Sequence 2220

25 MAKQPTALIIILDGFANRESEHGNVAVKQAHKPNFDRYIEKYPTTQIEASGLDVLPEQMG
 NSEVGHMNIAGRIVYQSLTRINKSIEDGEFFDNTVLNNAVKHVKDNGSALHVFGLLSDG
 GVHSHYKHLFAILELAKKQIDKVYVHAFDGRDQVQKSAKYIEETEDKFELGVGQFA
 SVSGRIYAMDRDKRWDREERAYNAIRNFEGPTFTSAKAGVEANYKNDVTDEFVEPFIVEG
 QNDGVNDGDAVIFYNFRPDRAAQLSEIFTNKAFDGFKVEQVDNLFYATFTKYNDNVDAEI
 30 VFEKVDLNNITIGEVADQDGLKQLRIATEKEYPHVTFYMSGGRNEEFEGERRRLIDSPKVA
 TYDLKPEMSAYEVKDALLEELDKGDLILLNFANPDMVGHSGMLEPTIKAI EAVDECLG
 EVVDKIIDMGHAIITADHGNSDQVLTDDQPMTHHTNPVPVIVTKEGVTLRETGRLGD
 LAPTLDDLNNVKQPSEMTGESLIKH*

35 Sequence 2221

Contig_0728_pos_3707_2403,
 is similar to (with p-value 0.0e+00)
 >gp:gp|AF065394|AF065394_1 Staphylococcus aureus enolase (en
 o) gene, complete cds. NID: g3152724.
 40 atgcaattattacagatgtttacgctcgcgaagtccttagactcacgtggttaaccraaca
 gttgaagttgaagtattaactgaaagtggtgctttcggacgtgcatttagtaccttctggt
 gcttctactggtgaacatgaagcagttgaattacgtgatggagataaatcacgttattta
 ggtaaaggtgtgactaaagcggtagaaaatgttaacgaaatgatcgaccagaaatcggt
 gaaggtgaattttcagtttttagatcaagtatctattgataaaatgatgattcaattagac
 45 ggtacacacaacaaaggttaaattaggtgcaaatgccatttttaggtgtttctattgcccga
 gctcgtgcagctgctgacttattaggtcaaccattatataaatatttaggtggattta
 ggtaaacaattgccagtacctatgatgaatattgttaatgggtggttctcactcagatgca
 ccaattgctttccaagagttcatgattttacctgtaggtgctgagtcattcaaagaatca
 ttacgttgggtgcagaaatctccataaccttaaatcaatcttaagtgaacgcggttta
 50 gaaactgcagtaggtgatgaaggtggtttcgacactagatttgaaggcactgaagacgct
 gtagaaactattattaaagctatcgaaaaagcaggatacaaacagggtgaagatgattc
 ttaggatttgactgtgcttcttctgaattctatgaaaatgggtgtttatgattacactaaa
 ttcaaggtgaaacacggtgctaaacgtagtgcagcagagcaagttgactacttagaagaa
 ttcgaattgttaaatatccaatcatcactattgaagatggtatggatgaaaacgattgggaa
 55 ggttggaacaattaaactgatcgtatcggtgataaagttcaattagttggtgatgattta
 ttcgtaactaaactgaaattttatctaaaggtatcgaacaaggtattggtaactcaatc
 ttaatcaaagttaaaccaaatcggtacattaactgaaacattcgatgctattgaaatggct
 caaaaagctggatatactgcggttgatctcaccgttctgggtgaaactgaagatactaca
 attgctgatatcgacgttgctacaaatgcaggccaaattaaaacaggttcattatctaga

actgaccgtattgctaaatacaatcaattattacgtattgaagatgaattatacgaaaca
gctaaatttgaaggaattaaatctttctacaatttagataaataa

Sequence 2222

5 MPIITDVYAREVLDSRGNPTVEVEVLTESGAFGRLVPSGASTGEHEAVELRDGDKSRYL
GKGVTKAVENVNEMIAPEIVEGEFSVLDQVSIKMMIQLDGTNKGKLGANAILGVSI
ARAAADLLGQPLYKYLGGFNGKQLPVPMMNIVNGGSHSDAPIAFQEFMILPVGAESFKES
LRWGAEIFHNLSILSERGLETAVGDEGGFAPRFEGTEDAVETIIKAIEKAGYKPGEDVF
10 LGFDCASSEFYENGVDYTKFEGEHGAKRSAEQVDYLEELIGKYPITIEDGMDENDWE
GWKQLTDRIGDKVQLVGGDLFVTNTEILSKGIEQGIGNSILIKVNQIGTLTETFDAIEMA
QKAGYTAVVSHRSGETEDTTIADIAVATNAGQIKTGSLSRTDRIAKYNQLLRIEDEL
YETAKFEGIKSFYNLDK*

Sequence 2223

15 Contig_0728_pos_2237_1779,
putative peptide of unknown function
atgacagactcaaagtctaaagaaataagaactggacgtttaattgcgataagttcatta
gtgttttgtattttactttatcatcacaccactttattgtatttagatgaatcaacagctaaa
tcaattttatcttttagctgggtcaaaaaacatcagatacagcagtgaaaaacatttttaaat
20 agtgaccgatacactggaattatgtatatatttttagcttacttagcaggtactgttgcttc
tggatcgccatccatattttatggtggtttatggttgcgatatatttctaatagcacta
tttacactcgtaaatctttactttatttcaaggtatttttagatgtaaaaaatgtactt
gcagttttaccaattttaattgtagtgttgatctataattctagcaatttatatgcta
gttggtttctattacacgtaaaagtactttcaatagatag

Sequence 2224

MTDSNAKEIRTGRLIAISSLVFCILLIIHHFIVLDESTAKSILSLAGQKTSDTAVKNILN
SDRYTGIMYILAYLAGTVAFWNRHPYLWWFMFAVYISNALFTLVNLYLFIQGILDVKNVL
AVLPILIVVIGSIIILAIYMLVVSITRKSTFNR*

Sequence 2225

Contig_0728_pos_1710_1147,
is similar to (with p-value 3.0e-60)
>pir:pir|A40585|A40585 DNA topoisomerase (ATP-hydrolyzing) (EC 5.99.1.3) chain B - Staphylococcus aureus >gp:gp|X71437|S
35 AGYRREC_2 S.aureus genes gyrB, gyrA and recF (partial). NID: g296393. >gp:gp|D10489|STAGYRABA_1 Staphylococcus aureus genes for DNA gyrase A and B, complete cds. NID: g540540.
atgcatacattaatcatcggttttattaattatagattgtattgcattagtgactgttgta
ttactccaagaaggtaaaaagtaattggactttcaggtgctattagtggcgctgaacaa
40 ttgtttggttaaacaaaaacaacgtggcgctgatttattccttgcatagattaacaatacgt
acgttattactttacattcttctatcggtttcatgagacctttaattgaagcgggctacgtt
tatattgctcagccgcctttatataaaactaacacaaggaaaaacaaaaatattatgtattt
aacgatagagaactagacaagttgaacaagaattaaacccgtcaccaaaatgggtcaatt
45 gcacgttacaaaaggtcttggtgaaatgaacgcagaccaattatgggaaacgactatgaat
cctgaacatcgctctatgttgcaagtgaagacttgaagatgcaattgatgcagaccaacaa
tttgaaatgttaatgggcgatgtagtagaaaatcgagacaatttatcgaagacaatgca
gtttatgccaacctagatttctag

Sequence 2226

50 MHFTLIVLIIIDCIALVTVLLQEKGKSNGLSGAISGGAEQLFGKQKQRGVDLFLHFLTIR
TLLL?FFYRFMRPLIEAGYVYIAQPLYKLTQGKQKYVFNDRDLKQLNELNPS?KWSI
ARYKGLGEMNADQLWETTMNPEHRSMQLQVRLEDAIDADQTFEMLMGDOVVENRRQFIEDNA
VYANLDF*

Sequence 2227

Contig_0728_pos_0_1110,
is similar to (with p-value 0.0e+00)
>sp:sp|P20831|GYRA_STAAU DNA GYRASE SUBUNIT A (EC 5.99.1.3).

atggctgaattacctcaatcaagaattaatgaacgaaatataaccagtgaatgcgtgaa
 tcattcttagactatgctatgagtggttatcggttctcgtgcattacctgatgttagagac
 ggattaaagccagtagcatcgctgattctttatgggttaaatagaacaaggtagatgagccc
 gataaaccttataagaaatctgcacgtatagtcggggatgtcatgggtaaatatccacct
 5 catgggtgattcttcaatttatgaagcaatggtaagaatggcccaagactttagttatcgt
 tatccactttagatgggtcaaggtaactttggctctatggatggtagcgggtgcagctgca
 atgctgttataccgaagcacgtatgactaaaataacattagaacttttacgtgatattaac
 aaagacacaattgattttattgacaactatgatggtaatgaaagagagccgctcagcttta
 cctgcacgtttccctaacttactagtaaatggcgccgaggaattgccgtaggtatggct
 10 acaaataattcctccccacaatttaactgaagttattgatgggtgctcagtttaagtaag
 aatccagacatcacaattaatgagctgatggaagacatacaaggtcctgattttcctaca
 gctgggttagttagtgaggaaaagtggtattcgtcgagcttatgaaacaggtcgtgggtca
 attcaaatgcgttctcgtgctgaaatagaagaacgtgggtggccgtcaacgtattgtc
 gtaacggaaatacctttccaagtcaataaagcgctgatgattgaaaaaatcgagagttta
 15 gtttagagataagaaaaatcgacggtattacagatttacgtgatgaaacaagtttgcgtaca
 ggtgtaagagtagttattgatgtacgtaaagatgcaaatgagtagttattttaaataat
 ttatataaaacaaacgccattacaacatcatttgggtggaatatgattgcttttagtgaat
 ggtagacctaaactaatcaatttaaagaagcacttatccattacttagaacacacaaaaa
 acagtggttagacgacgtactgaatataat

Sequence 2228

MAELPQSRINERNITSEMRESFLDYAMSVIVSRALEPDVRDGLKPVHRRILYGLNEQGMTP
 DKPYKKSARIVGDVMGKYHPHGSSIEAMVRMAQDFSRYRPLVDGQGNFGSMDGDGAAA
 MRYTEARMTKITLELLRDINKDTIDFIDNYDGNEREPSVLPARFPNLLVNGAAGIAGVMA
 25 TNIPPHNLTEVIDGVLSLSKNPIDITINELMEDIQGPDPFTAGLVLGKSGIRRAYETGRGS
 IQMRSRAETEEERGGGRQRIVVTEIPFQVKNKARMIEKIAELVRDKKIDGITDLRDETSRLT
 GVRVVIDVRKDANASVILNLYKQTPLOTSFGVNMIALVNGRPKLINLKEALIHYLEHQK
 TVVRRRTEYN

Sequence 2229

Contig_0730_pos_6751_0,
 putative peptide of unknown function
 atgcttaattctacaatcaaatatgacactaagaagaaaaaactacctaatttagtaaa
 ggtactaaaaagaaagacggtatattagatgttattagctctggtgtaaaaaatgatgtt
 35 aataaagtaaaagacattggtggtaaagcaagagacataggtggtactacgtttgacaaa
 gcaaaagacataggtacaaaagcacttgataaaagctaaagatgtgtctagcactgttatac
 aagggtattggagatgtttttgattatgtaggtcatcctatgaaattggtaataaaagtc
 tttgagaaagtgtgttttaacctagactttatgaaaaatgcaccattaccatttgattta
 atgacagctatgatttaagaaacttaaaaatggtattaaagacttctttaatgaaggttta
 40 gactctgcaggcggtggagatggttcttcgttactaaattcccaattactacggggtat
 taccctaaggtggtgctcctggttatagttttaatggtggtgctcactttggtattgac
 tatggcgctccataggtacaactatcaat

Sequence 2230

MLNSTIKYDTKKKKLPKFSKGTKKKDGILDVISSGVKNNDVNKVKDIGGKARDIGGTTFDK
 AKDIGTKALDKAKDVSSSTVIKIGIGDVFYVGHMKNLVNKFVKVGFNLDFMKNAPLPFDL
 MTAMIKKLKNGIKDFNEGLDSAGGGDGSSFTKFPITGYYPNGGAPGYSFNNGGAHFGID
 YGAPYGTTIN

Sequence 2231

Contig_0730_pos_6983_6669,
 putative peptide of unknown function
 gtgctagacacatcttagctttatcaagtgcttttgtagctatgtcttttgctttgtca
 aacgtagtaccacatgtctcttgcctttaccaccaatgtcttttactttattaacatca
 55 ttttttacaccagagctaataacatctaataaccgtctttcttttttagtacctttacta
 aatttaggtagtttttcttcttagtgctcatatttgattgtagaattaagcattatcaag
 gtcgcattcaatcaggtgagaagattaaataactaaataatattaggtgattcggtta
 tgcaaatattattag

Sequence 2232

VLDTSIALSSAFVPMFSALSNVPPMSLALPPMSFTLLTSFFTPELITSNIPSFLLVPLL
NLGSFFFLVSYLIVELSLIKVAFNQVRRNLNINILGDSLCKYY*

5 Sequence 2233

Contig_0730_pos_6435_6010,

is similar to (with p-value 3.0e-21)

>sp:sp|P42421|YXDJ_BACSU HYPOTHETICAL 26.6 KD SENSORY TRANSD
UNCTION PROTEIN IN IDH 3'REGION. >gp:gp|D14399|BACIOLO_11 Bac
illus subtilis 15 kb chromosome segment contains the iol ope
ron. NID: g709980. >gp:gp|Z99124|BSUB0021_70 Bacillus subtil
is complete genome (section 21 of 21): from 3999281 to 42148
14. NID: g2636442.

atcgatcaagtgatgatgagtggaacttggtgcagatgattatatgcaaaaaccattttat
15 acaaaacg!cttaattgctaagctacaagctatttatagacgcgtttatgaatttg:agtt
gaagaaaagagaaacgtaagttggcaagacgctactgtggatttatcaaaagatagatt
caaaaagatgataaaactatctttttgtctaaaacagagatgattattttagagatgta
atcaataaacgtaatacaatcgtagacacgagacactctcattactgctttgtgggatgat
gaagcttttgttagtgataatactttaacagttaatgttaataagataagaaaaaatta
20 tcagaaattgacatggatagtgcaattgaaacaaagttggtaaaggatacttagctcat
gaataa

Sequence 2234

MDQVMSMELGADDYMQPFYTNVLIQAIYRRVYEFVVEEKRTLSWQDATVDLSKDSI
25 QKDDKTIFLSKTEMIILEMLINKRNQIVTROTLITALWDDEAFVSDNTLTNVNRLRKKL
SEIDMSAIEETKVKGKGYLAHE*

Sequence 2235

Contig_0730_pos_5606_4977,

putative peptide of unknown function

atgaaattattgatagatcaagagaatgatgatcagcgtaagcgagcggttattatttgaa
tggtctcgtattaatgagatgtagataagcaattatatttaacaaggcgttgaaacacat
catcgtagatgtattttgattatatttcatataagagaatggttatagatgaaatacaa
gttactcgacatatcagtcaggcaaaaggataggttttgaaattagatttttaagacgaa
35 caaaaggtttatacagatgttaaatggtgccgtatgatgattaggcaagttctatctaac
tctttgaaatatagtgataattctacaataaatttaagtgggtataacatagaaggacac
gttggttttaaaattaaagactacggtcgtggaattagtaaaagagatttaccacgtata
tttgatagaggatttacttctacaacagaccgaacgatactgcgtcttctggtatggga
ttataccttgtacaaagcgtgaaagaacaacttgggattgaagttaaagttgattcaata
40 gtggggaaaggaacaacggtttatttccatttccacacaaatgaaatcattgagcgc
atgtctaaagtgacaagattgtcattttaa

Sequence 2236

MKLLIDQENDDQKRALLFEWSRINEMLDKQLYLTRLETHHRDMYFDYISLKRVIDEIQ
45 VTRHISQAKGIGFELDFKDEQKVYTDVKWCRMMIROVLSNSLKYSDNSTINLSGYNIEGH
VVLKIKDYGRGISKRDLPRIIFDRGFTSTDRNDTASSGMGLYLQSVKEQLGIEVKVDSI
VGKGTTFYFIFPQQNEIIERMSKVTRLSF*

Sequence 2237

Contig_0730_pos_4513_4070,

is similar to (with p-value 8.0e-32)

>sp:sp|P42423|YXDL_BACSU HYPOTHETICAL ABC TRANSPORTER ATP-BI
NDING PROTEIN IN IDH 3'REGION. >gp:gp|D14399|BACIOLO_13 Baci
15 llus subtilis 15 kb chromosome segment contains the iol oper
on. NID: g709980. >gp:gp|Z99124|BSUB0021_68 Bacillus subtili
s complete genome (section 21 of 21): from 3999281 to 421481
4. NID: g2636442. >gp:gp|D45912|D45912_2 Bacillus subtilis g
enome sequence between the iol and hut operon, partial and c
omplete cds. NID: g1408482.

- atgcttccactatcagttcaaaagttagataaacaataatgcatgaacgttatcaacgt
 atagtagaagctttgaacattagtgatattagtgataaatatccatcagagttgtcaggt
 ggacagcgtcaacgtacctcagcagcaagggcatttattaatttaccttcaattatatt
 gctgatgagcctacaggtgcttttagattctaaaagtacactagatttacttaagcgtctt
 5 aaatatatgaatgaggaatttaacacaaccatacttatggtgacacatgatccagtagcg
 gctagtttttcaaaccgtgttgatgttaaaggatggacaaatctttactgagttgtat
 caaggtgatgatgataaacaacgtttttataaagaaattataagaacgcaaagtgactt
 ggtggcatcaattatgagctttaa
- 10 Sequence 2238
 MLPLSVQKLDKQIMHERYQRIVEALNISDISDKYPSELGGQRQRTSAARAFINLPSIIF
 ADEPTGALDSKSTLDLLKRLKYMNEEFNTILMVTHDPVAASFNRVVMLKDGQIFTELY
 QGDDDKQTFYKEIIRTQSVLGGINYL*
- 15 Sequence 2239
 Contig_0730_pos_3876_3457,
 putative peptide of unknown function
 gtggctcttttgccttatgctaattttttattttgaaaagacgaggtcgtgaactatcg
 ttattacaaattattggtctaacaagaagatcatgaaaatgattatgttgagagcaa
 20 ttgatgacattttatgatgacaactattgttaggtatcatattgggaatctttggttcgaaa
 attttactcatgattgtattgcgattattaggaatcaacgtgagtggttctattatatt
 aattatcatgccattttagaaacgttattattaatagctgtgtcatatgtacttatagtc
 tttcaaagctatgtatatttacttaaacgttctattaaagagttagcgtctgatgtaaat
 aaaaaagagttcagtcatacacgcacaaacttggatgaagttgtattaggtttcttataa
 25
- Sequence 2240
 VVFLLYANFLFLKRRGRELSLLQIIGLTKKDIMKMIMLEQLMTFMMTTIVGIILGIFGSK
 ILLMIVLRLLGINVSIIIFNYHAILETLLLIASVYVLIVFQSYVYLLKRSIKELASDVN
 30 KKEFSHTRTTLGEVVLGFL*
- Sequence 2241
 Contig_0730_pos_3418_2012,
 is similar to (with p-value 2.0e-78)
 35 >sp:sp|Q02001|TRPE_LACLA ANTHRANILATE SYNTHASE COMPONENT I (EC
 4.1.3.27). >pir:pir|S35124|S35124 anthranilate synthase (EC
 4.1.3.27) alpha chain - Lactococcus lactis subsp. lactis
 >gp:gp|M87483|LACTRPOP_2 L. lactis trpE, trpG, trpD, trpF, t
 rpC, trpB trpA genes, complete cds. NID: gl49514.
 40 atggtatattgtatatacaaaaaggtgaatgctcaaattacgccagaagctttagcaaaatta
 aaacaaaaaagatcatttttgaaagtacaaatcaacagaaacttaaaggtaggtactcg
 atagtagtattcgatcattatggcaaaattacattagataattctcaacttttaattaag
 ttagacaatcattgtgaaatagttaagaatcaaccgtatcaacgacttaaggaatttgta
 gataaatattattttgaaatcaaagataaatatttaaaagatttaccttttatttcgggc
 45 tttatagggacatgtagctttgatttagtacgacatgaatttaaaaaattacaagatatt
 aaattagaagatcatcaaactcatgatgtccaattttatctagtggagatgtatttggt
 tttgatcattataaagatgaattatatattatcgcaagtaacttattttcttatagaaca
 aaagagagattaaaggaatctattgaacgtaaaattgaagatttaaaaaacatacatttt
 tcggttgaggatataaaattataaatccatccctcgacatataaccaccaatatatcagag
 50 caacaatttggttcaaactattagaatttttaaaaagaaaattactgaaggagatatgttt
 caagtagttccttcaagaatttatagttataaacaccattttcaacacaatttacatcaa
 ttaacttttcagttatatcaaaaatttaaaagcgacaaaatcctagtccatatatgtattat
 attataaagatgtaccgattgtaataggaagttctcctgaaagttttgtaaaggtaaaa
 gatgaaaagttttatagcaatcctatagctggaacaatttaaaagaggtcaaaaataaaaa
 55 gaagatgaaaataatgaaaagacattaatgaaagatgaaaaggaattgagtgaacatcgt
 atgctcgtagatttaggaagaaatgatattcatcgaataagtaaaacaggcacttcacaa
 attaccaactaatgacaatagaacgttatgaacatgtcatgcatatcgttagtgaagtt
 attggagaattaaaaccccatctatctcctatgagcgtcatcgcaagtttgctaccaacg
 ggtactgtctcaggtgcacctaaacttagagctatacagagaatatacgaatcttatcct

tataaaagaggtatctatagcgggtggtggttatatcaactgtaatcatcatttagat
 ttgcatggctatacgtaccatgattatcgatgaggaaaaagtcagtgctgaggcagga
 tgtggagtagtatatgattctattccagagaaagaacttgaagaaacaaaacttaaagct
 aaaagttattggaggttaactccatga

5

Sequence 2242

MDIVYKVNQAQITPEALAKLKQKKIIFESTNQQLKGRYSIVVFDHYGKITLDNSQLLIK
 LDNHCEIVKNQPYQRLKEFVDKYYFEIKDKYLKDLPFISGFIGTCSFDLVRHEFKKLQDI
 KLEDHQTHDVQFYLVEDVFDHYKDELYIIASNLFYRTKERLKEKIERKIEDLKNHIF
 10 SVEDINYKSI PRHITTNISEQQFVQTIRILKKKITEGDMFQVVPRIYSYKHHFQHNLHQ
 LTFQLYQNLKRQNPSPMYIINKDVPIVIGSSPESFVKVDGKVYTNPIAGTIKRGQNK
 EDENNEKTLMKDEKELSEHRMLVDLGRNDIHRISKTGTSQITKLMTIERYEHVMHIVSEV
 IGELKPHLSPMSVIAASLLPTGTVSGAPKLRAIQRIYESYPYKRGIYSGGVGYINCNNHLD
 FALAIRTMIIDEKVSVEAGCGVVYDSIPEKELEETKLKAKSLLEVTP*

15

Sequence 2243

Cortig_0730_pos_1446_451,
 is similar to (with p-value 8.0e-51)
 >sp:sp|P17170|TRPD_LACCA ANTHRANILATE PHOSPHORIBOSYLTRANSFER
 20 ASE (EC 2.4.2.18). >pir:pir|S42343|JS0340 anthranilate phosp
 horibosyltransferase (EC 2.4.2.18) - Lactobacillus casei >gp
 :gp|D00496|LBATRP_2 Lactobacillus casei DNA, trp operon (trp
 D, trpC, trpF, trpB, trpA), complete cds. NID: g216754.
 atgacccttcttgagaaaaattaaacaaaataaatctttatctaaaaaagatatgcaatca
 25 tttattgttacctggttattcaaatatagaaaccaatgtaaaggttgaattattgaaa
 gcttatacaaaataaagacatgggtcaatatgagctaactgatttagttgaatattttatc
 cagacaaaactatccaaaccaaccattttataataaagctatgtgtgtttgtggcacaggt
 ggagatcaatcaaatagctttaatatcttacaactgtagctttgtgtgtagcaagtgc
 ggagtgccagtcattaaacacggttaataaaagtattacttcacattcaggaagtacagat
 30 gtattacatgaaatgaatataaaaacaaaataaagaacgaagtagagcaacaattaaat
 ttgaaaggatttagcattcataagtgcgaactgattcttatccaatgatgaaaaagctcaa
 tcaattagaaaatcgattgcaacacctacaatttttaacttgattggaccattaattaat
 cctttcaaattaacttatcaagtgtgggggtatatgaagcttcacaacttgaaaata
 gcacaaaactataaaggatttaggtagaaaacgagcaattttaattcatggtgcgaatggg
 35 atggatgaggccacgctttctggtgaaaatatcatttatgaagttagcagcgaaagagca
 ttaaaaaaatatagtttaaaagcagaagaagtcggttttagcttatgcaataatgacacg
 ttgaaagtggttcacctaacaataaacaattgcattgaatatcctaagtgacag
 gatcactcaagtaaacgagatgtagttttgttaaatgctggaattgctttatatgttgc
 gagcaagtggaaagtatcaacatggcgtagagagagcgaaatatctcattgatacaggt
 40 atggcaatgaaacaatatttaaaaatgggaggttaa

Sequence 2244

MTLLEKIKQNKSLSKKDMQSFIVTLFDSNIETNVKVELLKAYTNKDMQYELTYLVEYFI
 QTNYPNQPFYNKAMCVCCTGGDQNSNFNISTTVAFVVASAGVPVIKHGKNSITSHSGSTD
 45 VLHEMNIKTNMNEVEQQLNLKGLAFISATDSYPMKKLQSIKRSIATPTIFNLIGPLIN
 PFKLTYQVMGVYEASQLENIAQTLKDLGRKRAILIHGANGMDEATLSGENIIEVSSERA
 LKKYSLKAEVGLAYANNDTLIGGSPQTNKQIALNILSGTDHSSKRDVLLNAGIALYVA
 EQVESIKHGVERAKYLIDTGMAMKQYLKMG*

Sequence 2245

Contig_0730_pos_447_70,
 putative peptide of unknown function
 atgactatttttaaatgaaattattgagtataaaaaaactttgcttgagcgtaaatactat
 gataaaaaacttgaaattttacaagataacggaatgttaagaggagaaagctgattgat
 55 tcacttaacttatgatagaacattatcagttattgctgaaataaaatcgaaaagcccatct
 gtacctcaattaccgcaacgtgatcttggtcaacaagttaaagattatcaaaaataagggt
 gctaactgctattttcaatattaactgatgaaaaatactttggcggtagttttgaacgatta
 aatcagttatcaaagataactgaacgtttaacctttgtacctcatcagcttgctcttcg
 acgcccattattcttatga

Sequence 2246

MTILNEIIIEYKKTLLERKYYDKKLEILQDNGNVKRRKLIDSLNYDRTL SVIAEIKSKSPS
 VPQLPQRDLVQQVKDYQKYGAN AISILTDEKYFGGSFERLNQLSKITERLTCTSSACSS
 5 TPIFL*

Sequence 2247

Contig_0732_pos_871_1170,
 putative peptide of unknown function

10 atggaccatattccacaagtagttatTTTTaataaaaaagacttatgtaacgaacagatg
 gatgtacctgtatctaaatctgcgcatgttttTgtatctagtctgtgatgaaatgataaa
 caaaaggTgaaaaatttagtaattcaagaaataaaaaatagtctcagcccatacgaagaa
 attgtagatagtgtgatgcagatagattatTTTTcttaacaacacacgcttggtact
 15 gaattaatatttgacgaaacacaagcatcttatcgatcaaaggatttaaaaaattataa

Sequence 2248

MDHIPQVVIFNKKDLQNEQMDVPVSKSAHV FVSSRDENDKQKVKNLVIQEIKNLSPLYEE
 20 IVDSADADRLYFLKQHTLVTELIFDETQASYRIKGFKKL*

Sequence 2249

Contig_0732_pos_3289_4629,
 is similar to (with p-value 0.0e+00)

>sp:sp|Q59812|GLNA_STAAU GLUTAMINE SYNTHETASE (EC 6.3.1.2) (GLUTAMATE--AMMONIA LIGASE) (GS). >gp:gp|X76490|SAGLNAR_3 S.a
 25 ureus (bb270) glnA and glnR genes. NID: gl134885.
 atgcaaaacgtagttttacaaaagatgatatttcgtaaaatttgctgaagaagaaaacgta
 agatattttaagattacaattcactgatatttttagggactattaaaaatgttgaaagtcca
 gtaagtcuattagaaaaagtattagataatgaaatgatgtttgatgggtcatctatgaa
 30 ggtttcggttcgtatcgaagaatcagatatgtatttacatcctgatttagatacttgggtt
 atcttcccttgactgctggacaaggaaaagttgcacgactaatctgtgatgtatttaaa
 acagatggtacaccatttgaaggatgccacgagctaactgaagcgtgtattaagaaga
 atggaagatatggccttactgattttaatctagggcctgaaccagaatttttcttattt
 aaattagacgaaaaagcggaacctacattagaattaaacgatgatgggtggttatttcgat
 35 ttagctcctacagatttaggtgaaaattgtcgccgtgacatcgttttagaattagaagat
 atgggctttgacattgaagcaagccaccatgaagtagcgccagggtcaacatgaaattgac
 tttaaatatgcagatgccgttacagcatgtgataatatccaaacatttaaaactagttggt
 aaaacaattgcacgtaagcataatttcatgcaacatttatgccaaaaccattatttggt
 gtaaacggtagtggtatgcacttcaacgtatcactatttaaaggaaaagagaatgcgttc
 40 tttgatcctgaaggtgatttacaattgactgatactgcataatcaatttacagctggtgtc
 cttaaaaaacgtagaggttactgcagtatgtaattccaattgtcaactcatataaacgt
 cttgtaccaggttacgaagcaccatgttatattgcatggagtggtaaaaaccgttcacct
 ttagtacgtgttccaacatctagaggtctatcaactcgtattgaagtacgctcagttgac
 cctgcagctaaccctgacatggcattagcagcaatcttagaagcagggttagatggaatt
 45 gagaataaacttgaggttccagaacctgtaaacaaaaatctacgaaatgaatcgtgaa
 gaacgagaagcgggttggtatccaagacttaccttcaactttatacactgcgttaaaagca
 atgcgtgaaaaataaatcaattaaaaacgcattaggtaatcatatttacaatcaatttatt
 aactcaaaatcgattgaatgggattactatagaactcaagtatccgaatgggaaacagaa
 cagtataattaagcaataactaa
 50

Sequence 2250

MPKRSFTKDDIRKFAEEENVRYLR LQFTDILGTIKNVEVPVSQLEKVLNEMMFDGSSIE
 GFVRIEESDMYLHPDLDTWVIFPWTAGQGKVARLICDVFKTDGTPFEGDPRANLKRVLRR
 MEDMGFTDFNLGPEPEFFLFKLDEKGEPTLELNDGGYFDLAPTDLGNCRRDIVLEMD
 55 MGFIDIEASHHEVAPGQHEIDFKYADAVTACDNIQTFLVVKTIARKHNLHATFMPKPLFG
 VNGSGMHFNVSFLFKKENAFFDPEGDLQLTDTAYQFTAGVLKNARGFTAVCNPIVNSYKR
 LVPGYEAPCYIAWSGKNRSPLVRVPTSRGLSTRIEVRSDPAANPYMALAAILEAGLDGI
 ENKLEVPEPVNQNIYEMNREEREAVGIQDL PSTLYTALKAMRENKSIKNALGNHIYNQFI
 NSKSI EW DYYRTQVSEWEREQYIKQY*

Sequence 2251

Contig_0732_pos_6496_6840,

putative peptide of unknown function

5 atgaataaagaccaaagaatcaatatgagttggaaaagctgttaaaagaaaatgaagaa
ctaaaagcagaaaaagctttatctcaaatgaagaatgagactcgttcaatgcttaatgag
tcagggttagaaaacttcgatgatcaaatgttaatatattagtaaaactgatgctgaa
aaaacaaggaaaaatgttgaatcatttactaacttacttaataaatggtaaaatcaaat
10 gttgaaaagcattaagacaagactcaccagtaagcactcaatcaataaaatgacaaaa
gatgaagaatcagattatcttgtttcaggagaaaaatgtctataa

Sequence 2252

MNKDQKNQYELEKLLKENEELKAEKALSQMKNETRSMNLNESGLENFDDQIVNILVNTDAE
KTRKNVESFTNLLNQMVKSNVEKALRQDSPVSTQSNKMTKDEESDYLVSGRKCL*

15

Sequence 2253

Contig_0732_pos_8090_8434,

putative peptide of unknown function

atgctaattattttaattttactagcattaagttttggtttaattcctattttctattaac
20 ttgcttatattaggtttaattccctttattttatggggagctatgggatgggctacacaa
gctcctcaacaacatatattattgaaaaaacatcctgaatatggaggctctgctgctcgt
ttaaatagttctattaattatttaggcagtgctatgggatcagcaatcggaggaattatt
ttatttaatgctaatagtacaaatgtactaataatagtgcttttaggaattactattatt
25 ggtattttattacaattactaaatttatccctagaaaaaaattaa

Sequence 2254

MLIILILLALSFLIPISINLPILGLIPFILWGAMGWATQAPQQHILLKKHPEYGGSAVA
LNSSINYLGSAAGGIIILFNANSTNVLIYSALGITIIGILLQLLNLSLEKN*

30

Sequence 2255

Contig_0732_pos_7495_7091,

putative peptide of unknown function

atggaaaaatcaacgcagacagacaaaacaaaactctgtgaacttaaaagcaaaacacaaaa
gatcaaaaataataacgcaaatgatgaagcagcttctccaactagcgaacaaaatgcagct
35 atagcacaagcaaaagtcatatgcaaatacattacctatctctaagaaaagttatacaaa
caattaacttcggaatacggagagaaaatatccggcagacatagcacagtatgctgttgac
catatcagtgtagattataaaatgaatgcactgagattagcaaaaagttacgtaaaaaat
ataaacattttctaatcaagcgttatatgatcaactcgtttcagaaaaatggagaaggattt
actcctgaagaagcacaatatgcaatgaatcatttagataggttaa
40

Sequence 2256

MEKSTQTDKQNSVNLKQNTKDQNNNANDEAASPTSEQNAAIAQAKSYANTLPISKKSLYK
QLTSEYGEKYPADIAQYAVDHSVDYKMNALRLAKSYVKNINISNQALYDQLVSENGEGF
TPEEAQYAMNHLDR*

45

Sequence 2257

Contig_0733_pos_5081_5728,

putative peptide of unknown function

gtgtctacttcccaattgattgtttcgaattccggacgagctaactcagggtttttctct
50 aattcagcaacagtggtgaatttagcgttagcacgttttaagatcgggtatttcccactt
gcagttgatactgaagttttttgtaccaattctgataagtcttgactgtcttaacttct
ttttcaggaataatatttaatatcctctgggtagttacgccaacgtcatcagatttaaca
ttgtcacgttttagcccttttgatttcatgtactgttcaaatgctagaatttcttcggtt
gtctctggattttgggttaatttagccatagaacgtttcgctccttctttttgtctttt
55 tcttttttaattcttctctgttggttcttctactttttcaatagtaggtgtttctggg
gtttcttcagggtttgtcatctgggtttgggtgcacatcaggtttttcttcatctgaagtt
ccttctgggtcatcatcagaaggtttgttctctgattcttctccagaattaccatctttg
ttatcttcaacttctgcaccttcatcttaggtggttcatcttgttttaggtgctgacgct
tcaatttcttttgaaagctgttcgagttcttcgtactctttcttttga

Sequence 2258

VSTSQLIVNSNGRANSNGFFSNSATVLNLALARFKIGYFPLAVDTEVFCTNSDKSWTVLTS
 FSGIYLISSGIVTPTSSDLTLRLAPDFDFMYCSNARISSFVSGFWFNLAIERFAPSFLSF
 5 SFFNSSSVGSSTFSIVGVSGVSSGLSSGFGASSGFSSEVPSSGSSSEGLFSDSSPELPSL
 LSSTSAPSSLGGSSCLGADASISFESCSSSSYSFF*

Sequence 2259

Contig_0733_pos_11496_10366,
 10 putative peptide of unknown function
 atgaattcgttgacatagcaggacgtattttcaaacagacgattcgagatgtaagaaca
 ttggcactgttacttattgcacctatattactattgtcgctactatattacatttttaca
 gttgccgataatacgaatggcgtaacagttggggttcacgatgtaccagattcattaatg
 actgaattacatgataaagatattcacgttaaacattataaaaaatgacaatgatataagt
 15 gataaaattaaagacgacaaattaacaggatttttgcacagtgatgggtcaaaaagtatca
 gtgacttatgctaacgataatcctacacaagcaggagaactaacagggtgcaaatcaaaaa
 tgggtaattgagtcataacatgaatgccatgaaagataataactaataaattgcatcaagcg
 ttaactaaaaatacaaaaaaatgcccggggatgggggagacacgcctcatcaagatatg
 gctaaccatataaactaacaacgcactatttatatgggtcatcagattctacgtatttt
 20 gatatgataaatcctatttttaattggattttttgtctttttctttacgttttttaattct
 ggcattggccttattaaaagagcgtacttctggcacattagaacgtttacttgcctctcca
 ataaaaagaagtgaattatttttgggttatgttttcggttatggtagtttttagcgttatc
 caaacaatagttgtcgtatttatatgcaatttatattctgcataatagacttagtaggttcg
 atatgggttcgtactattaacggcaatattaacagcgcttgcgctgtgacattcggtata
 25 ttattatctacctttgcttcctcagaattccaaatgattcaattttataaccattagtcata
 gtgccacaagtactatttgcaggcattataccaattgaatcaatgaataaaggattacaa
 tacttttcacatatcatgccgttattctataccggccaaacgatgcaaaatattatgatc
 aagggttatggattcaacgatatttacatttatttaattgtgtatttcgcatttttcatt
 ttcttattgatttttaaatattatagggcatgaaaagatatagaaaagtttag
 30

Sequence 2260

MNSLHIAGRIFKQTIIRDVRTLALLLIAPILLLSLLYYIFTVADNTNGVTVGVDV: DSLM
 TELHDKDIHVHKYKNDNDISDKIKDDKLTGFLHSDGQKVSVTYANDNPTQAGELTGANQK
 WLMShNMAMKNDNTNKLHQAALTKIQKMPGDGGDTPHQDMAKPYKLTHYLYGSSDSTYF
 35 DMINPILIGFFVFFFTFLISGIGLLKERTSGTLERLLASPIKRSEIIFGYVFGYGSFSVI
 QTIVVLYAIYILHIDLVGSIWVLLTAILTALVAVTFGILLSTFASSEFQMIQFIPLVI
 VPQVLFAGIPIESMNKGLQYFSHIMPLFYTGQTMQNMIMIKGYGFNDIYIYLIVLFAFFI
 FLLILNIIGMKRYRKV*

Sequence 2261

Contig_0733_pos_10356_9706,
 40 putative peptide of unknown function
 atgaaccaagatattaagtcatttagttgaaaccattgtgcctcaacttgaatatttaagc
 gataaacaagacgtgtcatagaaagtgtcatttgcattattcagtgaacaaggatttgat
 45 aaaacgagtagtaaaagaattgcgcagcgtgcaaatgtcgcagaaggaacggtatttaag
 cagtttaaaagtaaaagaatgttattatacgcaggatttaattccaatttttaagagatcat
 atcgcacctgtagctgttaaaacaatttacagatgaattaaacgaagtaaccattttgat
 gcatttataaaatttttagtaaaatagatctaaatttttatgacaatagacgtatt
 cttaaagtcatttaaatgaagctattactaatgaagattttcaaaatatattagttaat
 50 attttcacccataaattaacgagtaaaattaaaagataaaaattgaatggtttatcgataat
 ggtgacatgcgcaatgttaaacctgagttttttatacgtacggctcgtcgcaaaaattta
 aatttaaatatcccaataatagtttaataatgactataactaagggtgaaaactatcagcag
 tttgcgttatttcgtaaaagagggttatataggatgtttaagcgagaatag

Sequence 2262

MNQDIKSLVETIVPQLEYLSKQRRVIESAIALFSEQGFDKTSTKEIAQRANVAEGTVFK
 QFKSKRMLLYAGLIPILRDHIAPVAVKQFTDELNEVTHFDAFINLFVENRSKFIYDNRI
 LKVILNEAITNEDFQNILVNIFTHKLTSKLKDKIEWFIDNGDMRNVKPEFFIRTVVAQIL
 55 NLNIPIIIVNNDYTKGENYQQFALFVKEGLYRMFKRE*

- Sequence 2263
Contig_0733_pos_4489_3998,
putative peptide of unknown function
- 5 atgccactttttgtgacagattattttaaataaaaaggaaaagatatgtaggctatttagca
agtttatgtgcaggtttagaagtgccgtttatggttatattaggtatattatcagctaaa
ttgccaactcgaactttattgataattggttctgtatttggcgcgcatttactttagt
attggtgtatttataaaattttcatatgatgcttgggttgacaaatatgtttagccatcttt
ctagcaatattactaggctcttgggattagttattttcaagatatcttccctgattttcca
10 ggttatgcatcaacactttttgccaatgccatggtaataggacaacttgggtggaatttg
ctagggtggtgtgatgagtcattgggttgggttggaaaatgtattctttgtctcagcaagt
tctatcttcgtaggatgatactcatcttgtttacgaaaagatcaaaaaattacaatagaa
gatgtggagtag
- 15 Sequence 2264
MPLFVTDYLNKERYVGYLASLCAGLEVPMVILGILSAKLPTRTLIIIGSVFGGAFYFS
IGVFNKNEHMLVGQICLAI FLAILLGLGISYFQDILPDPFGYASTLFANAMVIGQLGGNL
LGGVMSHWVGLENVFFVSASSIFVGMILILFTKDQKITIEDVE*
- 20 Sequence 2265
Contig_0733_pos_3991_3503,
putative peptide of unknown function
- 25 atgacagcgatattatggattttaattattatagcctttgcgtagcatttattggttta
attaaaccgggtgataccatcacttttaattgttatggattggtttttaattatcaattt
ggttttcatgaggggaagattatcgtggattttttatgttgcaatgattatctttaccatt
atgatattagtagccgattttgtgatgaataaataatttcgtcaatcgctttggaggaagt
aaaatagggtgaatacacagcgctcataggtgtgattgttaggttgtttcgttttccctccc
tttggattatcattatcttttgggtgctgttgcattgttgaattggttcaagggttt
aactttcaacaagctataaagggtgagttttggctcagtgattgcatttttagcgagtaca
30 attgctcaagggtctaataatgattgtaatggttatttgggtcttttttagatgtctttcta
ataaattaa
- Sequence 2266
MTAILWILIIIAFALAFI GLIKPVIPSLMLWIGFLIYQFGFHEGRLSWIFYVAMIIFTI
35 MILVADFVMNKYFVNRFGGSKI GEY TALIGVIVGCFVFPFPGIIIIIPFVAVFIVELVQGF
NFQQAIKVSFGSVIAFLASTIAQGLIMIVMWIFFLDVFLIN*
- Sequence 2267
Contig_0733_pos_2319_1897,
putative peptide of unknown function
- 40 atgagaaaatggttaaccttactattaattacaacattggtgttaactgcatgttggtaaa
agtaacgaaaaagcttctttagaaaaaagcattgatcagttgaaaaaagaaaaataaggat
ttaaaaaacagaagaaaaagttacaagagcaaaaggataagcttaaacacaaaacaggat
agtctccaagaagatgtaaatgacttgccctgctaaaagcacatcccagataagaaaaat
45 aaagataatcatgatgcaaaagaaaagcttctcgataatcaatcgacatctgctaataat
gatgatcaaaactaacaataaaaaagcaatcaagatgaacatgacagtcattcctctaaa
ccacatacacagcagaagccctcacagaatgatagaaaaataatcatcgacaagaacga
tag
- 50 Sequence 2268
MRKWLTLTLLITTLVLTACGKSNEKASLEKSIDQLKKENKDLKKQKKKLQE QDKLKHKQD
SLQEDVNDLPAKSTSRDKKNKDNHDAKEKSSDNQSTS ANHDDQTNKIKSNQDEHDSQSSK
PHTQQKPSQNDKRNNHRQER*
- 55 Sequence 2269
Contig_0733_pos_765_244,
putative peptide of unknown function
- atgtcaaaaatcttaaacacacaattaactggtatttttaactcggttgaaaaacaagag
ttggatattcaaatggcagctcaatgtctcattcaagcaattggtggagaaggacatgtc

tatatcaaaggctacgatgattttaaattctatgagtcattcatattacaaagccatgaa
 aaattagcgtctagcttaccacttgaagatttacaaaattttaacgatatagatacaaca
 gatagggtagtggtattttcaccatactacacttcggaagtgaaagtgatgtacttcaa
 cttattgatttagatgtcgatttagtgcttatttgtaataaccctaaacgagatgatttt
 5 cctaatacatttaattcattatgttaatttatcaacacctaggccattgtttacacagaa
 gattatgataaaatcattcaaccacatccgatggccttaaattatatttattatgatatt
 tatactcaaagtattgagatgactagagacctagatttatag

Sequence 2270

10 MSKILNTQLTGIFNRLEKQELDIQMAAQCLIQAIGGEGHVYIKGYDDLKFYESFILQSHE
 KLASSLPLEDLQNFNDIDTTDRVLLFSPYYTSEVESDVLQLIDLVDLVLICNNPKRDDF
 PNHLIHVNLSTPRPIVYTEDYDKIIQPHMALNYIYDIYTQMIEMTRDL*^{*}

Sequence 2271

15 Contig_0734_pos_2644_3024,
 putative peptide of unknown function
 atgcaccagcttgtccatactgacgataatgcaataaatgtagatgtattaccaccacaa
 caagctgacggtaaacgactaatccagaacaattatttgctgcaggttacgcttcagtc
 ttaattggtgcatttgatttaatttttaaaacaaaataaagtgcgcgatgctgaacctgaa
 20 gtaacgttaacggtacgcttggagacgatccagatgccgaaagcccaaaacttagcggt
 gatattcatgcaaaagttaaaaatgttttatcacaagaagatgctgaaaaatatttaca
 gatgcgcacgacttttgcgtattcaaaagctacacgtggcaatatcgatgtaaaactta
 aatgttgaagtagtagaataa

Sequence 2272

25 MHQLVHTDDNAINVDLPPQQADGKATNPEQLFAAGYASCFNGAFDLILKQNKVRDAEPE
 VTLTVRLEDDPDAESPKLSVDIHAKVKNVLSQEDAKEYLQDAHDFCPYSKATRGNIDVNL
 NVEVVE*^{*}

Sequence 2273

30 Contig_0734_pos_3557_4369,
 is similar to (with p-value 1.0e-96)
 >gp:gp|Y17554|BLY17554_1 Bacillus licheniformis arcA, arcB,
 arcC and arcD genes. NID: g3687415.
 35 atggatgataaatccggttttatcttgaccctaatgcctaattctatattttcacacgtgat
 cctcaagcttcaattggtagaggtatgacagtaaatcgatgttttggagagcgagacgc
 agagaatcgattttcatattcatatattttaaacatcatcctagattttaaagatgagaat
 attcctttatgggtggatcggtgactgtccgttcaacatcgaaggtggagacgaactgggtg
 ttatctaagatgtacttgcaatagggatatetgaacgtacttctgcacaagcaattgaa
 40 cgtttagcacgacgtatttttaaagatccggttatctacttttaaaaaggtggtggcgatt
 gagattccaactagtcgaacatttatgcacttagatactgtttgtacaatgattgattac
 gacaaattcactacacattcagcaattcttaaatcagaaggaaacatgaatatctttatt
 atcgaatatgatgataaagctgaagatatcaaaatccaacattctagtcattcttaacaa
 acattagaagaagtgcctgatgttgatgaaatcacattaataccaactggaaatggtgat
 45 atcatcgacggtgctcggaacaatggaatgatggttcgaatactttatgcatacgctccc
 ggtgtggttgtaacttatgatcgtaattatgtttctaataattgttacgtgagcatggt
 atcaaagttattgaaattcctggaagtgaactgtacgtggctgaggaggccctcgatgt
 atgagtcacaccttaataagagaagatctatag

Sequence 2274

50 MDDKYPFYLDPMPLNYFTRDPQASIGRGMVNRMFWRARRRESIFISYILKHHPRFKDEN
 IPLWVDRDCPFNIEGGDELVLISKDVLAIIGISERTSAQAIERLARRIFKDPSTFKKVVAI
 EIPTSRFTMHLDTVCTMIDYDKFTTHSAILKSEGNMNIIFIYDDKAEDIKIQHSSHLKQ
 TLEEVLVDVEITLIPTGNGDIIDGAREQWNDGSNTLCIRPGVVVYDRNYVSNQLREHG
 55 IKVIEIPGSELVRGGPRCMSQPLIEDL*^{*}

Sequence 2275

Contig_0734_pos_4456_5061,
 is similar to (with p-value 2.0e-70)

>gp:gp|AJ001330|LSAJ1330_2 Lactobacillus sake DNA encoding the arginine-deiminase pathway genes. NID: g2764610.
 atgaaagggcgcatcccatgaaaaacatcaagaaaccctttgatttaaaggtaagtca
 ttgctaaaagagtatgatcttacaggtgaagaatttgaaggtctaatacgattttgctatg
 5 acattaaaaaaatataaacaacaaggcacaccacatcgatatttagagggtaagaatatt
 gctttactcttcgaaaagacatctactcggacgcgtgccgcatttacagttgcatctatt
 gatctaggtgcacacctgaatttttagggaaaaatgatattcaattaggaaaaaaagaa
 tctgttgaggatactgctaaaagttttaggcagaatgtttgatggaattgaatttagaggt
 ttttcccaaaaaactgttgaacaattggccgaattctctggagtaccagtatggaatggg
 10 ttaactgatgattggcatcctacacaaatgttagctgattatatgacaattaaagaaaat
 tttggatatttaaaaggcatcaacctaacttatgtaggaaacggacgtaataatgttgca
 cattcgcttatggtggcggtgctttcaatattcaatctgatatgggagggccacgctggag
 gattag

15 Sequence 2276
 MKAHPMKNIKKPFDLKGKSLKEYDLTGEEFGLIDFAMTLKKYKQQTTPHRYLEGKNI
 ALLFEKTSTRTRAAFTVASIDLGAHPEFLGKNDIQLGKKESVEDTAKVLGRMFDDGIEFRG
 FSQKTVEQLAEFSGVPVWNGLTDDWHPTQMLADYMTIKENFGYLGKINLTYVGNRNNVA
 HSLMVAVLSIFNLIWEATLED*

20 Sequence 2277
 Contig_0734_pos_5331_6341,
 is similar to (with p-value 4.0e-96)
 >gp:gp|Y17554|BLY17554_3 Bacillus licheniformis arcA, arcB,
 25 arcC and arcD genes. NID: g3687415.
 gtggcatcaatattattgtggagtggtcactttttaataattaaaaggagttgagaaggct
 gctttaatcaatagttatgtcactattacaaaattaattcctatcttattagtattata
 tgtatgattgtagccttttaatttaatacattcaagataggtttctttggaatggatgga
 tatggatcattatcattttcgtctaatacagatgtcacagttaaaagtagcatgtta
 30 gtgacagtttgggtatttattggtattgaaggtgctgttattctccggaagagctaaa
 aataaaaaagatgttggaaactgccactgttatcgacttatttcagtattgctcatttac
 ttcttgctgactgtatttagcacaaggtattgtaattcaaaatcatatttctaaacttgag
 gcaccatcaatggcacaatttttagcttatattgttgggtgattggggagcaacatttgtc
 aatattggtcttattatttcagtattaggtgcatggctaggttgacattacttgccgga
 35 gagttgccttttattgttagctaaagatgggtttgttcttaagtgggttgcaaaagaaaat
 aaaaatggggcacctatgaacgccttattcattactaatgtgttagttcaaatattcctt
 attagtatgctgtttaccaaagtgcttatcattttgcattttctctcgcggaagtgtc
 atactttatccatacatgttcagtgctgttttatcaggtgaaatatacaatagaacacaag
 ttaactgcaacgcctaacaatggattataggaattctagcatctatttatgcaatttgg
 40 ctggtatagcgtcaggtatagattacttgttacttaccatgttgctctatattccgggg
 attatcgctctatgttgttggccagaaaataatcaaaagcgacttacacaatttgactat
 atttcttcagtccttatcggtatttttagcattgatagggttattacgatga

45 Sequence 2278
 VASILLWSVHFLILKGVETAALINSIVTITKLIPILLVVICMIVAFNENTFKIGFFGMDG
 YGSLSFHFANTMSQVKSTMLVTVWVFIGIEGAVVFSGRAKNKKDVGATVIGLISVLLIY
 FLLTVLAQGIVIQNHISKLEAPSMQILAYIVGDWGATFVNIGLIISVLGAWLGWTLLAG
 ELPFIVAKDGLFPKWFAKENKNGAPMNALFITNVLVQIFLISMLFTKSAYHFAFSLAASA
 ILYPYMFSAFYQVKYTIHKLATATPKQWIIIGILASIYAIWL VYASGIDYLLLTMLLYIPG
 50 IIVYVVVVQKNNQKRLTQFDYIFFSLIVILALIGLLR*

Sequence 2279
 Contig_0734_pos_2115_1636,
 putative peptide of unknown function
 55 gtgccagatcatatagagaaagttagtggtcgtagtaaatcctcaaatgtccaccataaag
 agaataattaatcaactgatattaacacaatccaattacatggaaatgaaagcattcaa
 ttaattagaaatattaagaaacttaattcaaaaaataagaatcataaaagcaattccagca
 acaagaaatttaataataaacattcaaaagtataaagatgagatagacatgtttattata
 gatacaccatcaatcacatacggaggagcaggtcaaagttttgactggaaattattaaaa

aaaataaagggcggttgattttctcattgcgggtggtttggattttgaaaagataaaacga
 ttagaaatatattcatttggacaatgtggttatgacatctcaactggcattgagtcacat
 aatgaaaaagattttaataagatgactcgaatattaaaatttttgaaaggagacgaatga

5

Sequence 2280

VPDHIEKVVVVVNPQMSTIKRIINQTDINTIQLHGNEISQLIRNIKKLNSKIRIIKAIPA
 TRNLNNNIQKYKDEIDMFIIDTPSITYGGTGQSFDWKLKKIKGVDFLIAGGLDFEIKIR
 LEIYSFGQCGYDISTGIESHNEKDFNMTRILKFLKGDE*

10

Sequence 2281

Contig_0734_pos_1632_424,

is similar to (with p-value 0.0e+00)

>sp:sp|Q01998|TRPB_LACLA TRYPTOPHAN SYNTHASE BETA CHAIN (EC
 4.2.1.20). >pir:pir|S35129|S35129 tryptophan synthase (EC 4.
 2.1.20) beta chain - Lactococcus lactis subsp. lactis >gp:gp
 |M57193|LACTRPOP_7 L. lactis trpE, trpG, trpD, trpF, trpC, t
 rpB trpA genes, complete cds. NID: g149514.

atgaaaattcaaacagaagtagatgaattgggctttttcgggtgaatatggtggccaatat
 gtacctgaaacattgatgccagctattattgaacttaaaaaagcatatgaggacgcgaaa
 tcagatactcacttcaagaagaatttaattattatttaagtgaatatgttggtagagaa
 acgcctttaacatttgctgaatcacacaaaattgttaggtgggtgccaataatatctt
 aaaagagaagacttaaatcacactggtgctcataaaattaataacgcgataggacaggca
 ctattagctaaaaggatggggaaaactaaattagtagccgaaacaggtgctggtcaacat
 25 ggtgtagcaagtgccaccatcgctgctttattcgatatggatcttattgtttcatggga
 agtgaagatatcaaacgtcaacaacttaacgtatttagaatggaattgctaggagctaaa
 gtagtgtctgtgcagatgggcaaggaaacactatcagatgctgtaataaaagctttgcaa
 tattgggtgaatcatgtcgaggatacacattattttattaggctcagcgttgggacctgat
 ccgtttccaactatggtcagagattttcagagtgtgattggtaataaaagccaa
 30 attttaagtaaaagaaggacgacttccagatgcgttagtcgctgtgttggtggaggatcc
 aattcaataggtacgttctatccattatatacaagatgatgttaaattataggggtagaa
 gctgcgggaaaaggagtcatacgcataatcatgcttagctatagggaaggtaaacca
 ggtgtattacatggttccaaaatgtaccttattcaaaatgatgatggacaaattgaattg
 gcacactctatcagcgggactagattatccaggtattggacctgaacattcgtattat
 35 aatgatattggtcgtgtatcatatgtaagtgtctacagataatgaagctatggaagcactt
 ataacattctcaaaagttgaaggtatcattccagcaattgaaagtcacatgcattgagt
 tatgttgaaaattagcgccaaatatggatgaaaaagaaattattgttgtgactatttca
 ggtcgtggagataaagatatggaacaattaacaatacaagaaaacggtggtgaacaa
 aatgagtaa

40

Sequence 2282

MKIQTEVDELGFFGEYGGQYVPETLMPAIIELKKAYEDAKSDTHFKKEFNYYLSEYVGRE
 TPLTFAESYTKLLGGAKIYLRDLNHTGAHKINNAIGQALLAKRMGKTKLVAETGAGQH
 GVASATIAALFDMDLIVFMGSEDIKRQQLNVFRMELLGAKVVSVDGQGLSDAVNKALQ
 45 YWNVHVEDTHYLLGSALGPDFFPTMVRDFQSVIGNEIKSQILSKEGRLPDALVACVGGGS
 NSIGTFYFPFIQDDVKLYGVEAAGKGSHTNNHALAIGKKGKPGVLHGSKMYLIQNDGQIEL
 AHSISAGLDYPGIGPEHSYNDIGRVSYSVATDNEAMEALITFSKVEGIIPAIESAHALS
 YVEKLAPNMDEKEIIVVTISGRGDKDMETIKQYKENGGEQNE*

50

Sequence 2283

Contig_0734_pos_0_404,

is similar to (with p-value 8.0e-20)

>sp:sp|P17166|TRPA_LACCA TRYPTOPHAN SYNTHASE ALPHA CHAIN (EC
 4.2.1.20). >pir:pir|S42347|JS0344 tryptophan synthase (EC 4.
 2.1.20) alpha chain - Lactobacillus casei >gp:gp|D00496|LBA
 55 TRF 6 Lactobacillus casei DNA, trp operon (trpD, trpC, trpF,
 trpE, trpA), complete cds. NID: g216754.

atgggtgatttaaatatttcatcatttaaaaacattaactgagaatggagcagacatt
 gttgaaattggtgtgccattttctgacctgttgcatgagatggacctataatcatgaaagca

5 gggcgcaacgctattgacgaggggtcaaacattaaattcatttttgatgaattaataaaa
 aataaaaaatactatttcatctaagtagtattatgacttattataatattctaagtgc
 tatggagaagaattatttttggataagtgatgaagctgggtgtttatgggttaattatt
 ccagatttaccttacgaacttacaaaaaggtttaaaaaagatttttatcatcattctgtt
 aaaataatatcgtaattgccatgaccgcaagtgatgctaggat

Sequence 2284

MGDLNFIHHLKTLTENGADIVEIGVPFSDPVADGPIIMKAGRNAIDEGSNIKFI FDELIK
 NKNTSSKYVLMTYYNILSAYGEELFLDKCDEAGVYGLIIPDLPYELTKKFKKDFHHSV
 10 KIISLIAMTASDARX

Sequence 2285

Contig_0735_pos_1041_2267,
 is similar to (with p-value 3.0e-89)
 15 >gp:gp|AF024571|AF024571.1 Staphylococcus aureus high affini
 ty proline permease (putP) gene, complete cds. NID: g2565310

atgagtttagtatatgacaaacttactcaagatcaaccgtatcattcgtgggttaattatt
 gttgagcatttcttaccttctgatagtcattttagatattggttgcggtactggc
 20 aacttaacacaataactaactgactaggtgaagtcactggatggatattagtgtagat
 atgttatcaatagctagacaaaaaacaatcaagtgaagtgatcgaaaggaatatgact
 cactttaatttgacaaaaaatttaatatgattacaatatgttgattcactgaattat
 ttagaaacattaaatgacgtaaaaaatgacattcgaaagagtgtatcaacatttaataaaa
 aatgggtgtttttatttttgatgtacatactgttcataaaatgaaaacattatttaataat
 25 aaaagttatattgatgaatctgataatgtttttgtagggtgggatgcaatatgtggggat
 gaaccatttagtaagatttatgtctatcaaatcacataaattactacaaaagccagacgt
 ctgggaataagttggatggcagtcggtctattaggagcgattgggtgtaggattaacagga
 atttcatttatatctgaaagacatatataatcagaagatcctgaaacactatttat.tgtg
 atgagtc aaatattatttcatccgctttaggtggatttttattagcagccatccctgct
 30 gcaataatgagtactatctcttcacaattactagtaacatcaagttctttaactgaagat
 ttctataaactaatcagaggttcagataaagcatcatcacacaaaaagagtttggtttg
 attggacgcttatcagttctacttgttgcgtagttgctattacgattgcttggcatcca
 aacgatacaataactaaatttagttggttaattgcttgggctgggttttgagctgcatttagt
 ccttttagtactctactcttttatattggaaagatttaacacgtgcaggagctattagcgg
 35 atggtagctgggtgctgtggttgttattgtttggatttcttgataaaacccttggctaca
 atcaatgcattctttggtatgtatgaaatcattccaggtttcataattagcgtattgatt
 acctacatcgtaagtaaatcaaaaaaacctgatgattatgttattgaaaatcttaata
 aaagttaaacacatcgtaaaagaataa

40 Sequence 2286

MSLVYDKLTQDQPHYHSWFNIVEHFLPSDSHLLDIGCGTGNLTQLLTS LG EVTGM DISVD
 MLSIARQKTNQVKWIEGNMTHFNLNKKFNMITIFCDSLNYLET LNDVKMTFERVYQHLNK
 NGVFI F DVHTVHKMKTLFNNKSYIDESDNV FVGWDAICGDEPLVRFMSIKSHKLLPKARR
 LGISWMAVGLLGATGVLTGISFISERHIKSEDPETLFIVMSQILFHP LVGGFLLAAILA
 45 AIMSTISSQLLV TSSSLTEDFYKLIRGSDKASSHQEFVLIGRLSVLLVAIVAITIAWHP
 NDTILNLVGNWAGFGAAFSPLVLYSLYWKDLTRAGAI SGMVAGAVVVIWISWIKPLAT
 INAFFGMYEII PGFIISVLITYIIVSKLTKKPPDDYVIENLNKVKHIVKE*

Sequence 2287

50 Contig_0735_pos_4945_3824,
 is similar to (with p-value 0.0e+00)
 >gp:gp|AJ011676|BST011676_1 Bacillus stearothermophilus lig
 gene. NID: g3688228.
 atgggttatacgcaaaaatctccaagatggcgattgcttataaatttccagctgaagaa
 55 gttattacaaaattattggatattgagctaagattggcgctacgggtgtgtgacacca
 actgcaattctagaacctgtaaaagtagctggtagctacagtttcaagagcctcacttcat
 aatgaagatttaatacatgaaagagatatcgtatcggagatagtggttattaaaaaa
 gccggggacatcatccctgaagttgtaaaagatttttagatagacgacctaacgaatcg
 gaaatttatcatatgccaacacattgtccttagttgtggacatgaattagttcgtattgaa

ggagaagttgctttacgttggtattaatccaaaatgtcaggcacagcttattgaaggactt
 atacatttcgtttcaagacaagcgatgaatatagatggtttaggtactaaaattattcat
 cagctatacgaataacagtttaatacaagatgtcgagatattttctatttgaaagaagaa
 gatttattaccattagagcgaatgggaaagaagaaagttgataatcttttattagcgata
 5 gaaaaatctaaagaacagtcatttagagcatttattatttggacttggtattagacattta
 ggtgtaaaagctagtcaggtacttgctgagcgatatgaaacgatggatcaactttttaaa
 gtaactgaaagtgaaatgaattgaaattcaagatattggagataaacttgacaaatctgtt
 gtaacatatctcgaaaatagtgtatctgttcatttaattgaaaaattaagtaataaaaaat
 gttaatatgtcttataaaggaattaaaacaactgaaatcgaggtcatcctgatttttagt
 10 gggaaaacaattgtatttaacagggaaactcgagcaaatgacgagaaatgaagcatctgaa
 tgggtgaaaatgcaaggtgctaaagttacaagcagcgtgactaaaagtactgatattgtc
 atagctggagcagatgcaggggtctaaattagccaaagctgagaagtatggtactgaaatt
 tggactgaagcagcatttattgaaaaacaaaatggaatctaa

15 Sequence 2288

MGYTQKSPRWAIAYKFPAEEVITKLLDIELSIGRTGVVTPTAILEPVKVAGTTVSRASLH
 NEDLIHERDIRIGDSVVIKKAGDIIPEVVKSIILDRRPNSEIYHMPHCPSCGHELVRIE
 GEVALRCINPKCQAQLIEGLIHVSRQAMNIDGLGTKIIHQLYENQLIKDVADIFYLKEE
 DLLPLERMGKKKVDNLLLAIEKSKEQSLEHLLFGLGIRHLGVKASQVLAERYETMDQLFK
 20 VTSEELFIQDIGDKLAQSVVTYLENSDIRSLIEKLSNKNVNMYSYKGIKTETIEGHPDFS
 GKTIVLTGKLEQMTRNEASEWLKMQGAKVTSSVTKSTDIVIAGADAGSKLAKAEKYGTEI
 WTEAAFIKQNGI*

Sequence 2289

25 Contig_0735_pos_3414_2596,
 is similar to (with p-value 1.0e-49)
 >gp:gp|Z99107|BSUB0004_111 Bacillus subtilis complete genome
 (section 4 of 21): from 600701 to 813890. NID: g2632866. >g
 p:gp|Y15254|BSYERABCD_7 Bacillus subtilis 13kB DNA fragment,
 30 from yerA to sapB gene. NID: g2577959.
 atgagcgaaaaagaaaaagaaaaagcaaaaatgctaatagagaatcttggactcaatccatct
 cacaatggtgaaacagatgaagagaaaaatagctaaaaattctccagcctatctttcaaat
 atactcgagcaggatttttatggaaatagtgtattctaaaggtaaaaatataaaagggatg
 acaattgggttttagctatgaatagtgtttattattacaaaaaagagaaagatggcgaaaca
 35 ttttagtaagattttatctgataaagagattgaaaagcaaggtaaacagatggctagtga
 atgctttctcgtttacgtgagaatagtgtattgaaagataattcctattcattttgctatc
 tataaacaatcaagtcaagattccattacaccaggtgaatttatagttggtactacggtt
 gaagagggtaaaactaaaattaactcatgggataatattaatgaaaaagcagccttaatt
 ccttcgtcaactgcagctgattatgatgaaacggtgaataataactttaaacagtttaaat
 40 gataatttcgcaactcgtatttttcaaaactcacacaagcagttggtaagggttaaatcgt
 aataaaaaagctaacaacttacagttgattttgcctatagattattacggacaggcagaa
 acgataggtattacacaatatgttacagagcaagccgaaaaatattttgataaaactagat
 gagtatgaaattagaatcaaagatggaataactccacgtgctctcattagtaaaactaaa
 gacgataaagaaccacaagttcatatctatcataattag

45 Sequence 2290

MSEKEKKSKNANENLGLNPSHNGETDEEKIAKNPAYLSNILEQDFYGNSSDKGKNIKGM
 TIGLAMNSVYYYYKKEKDGETFSKDLSDKEIEKQKQMASEMLSRLENLKDIPHF
 YKQSSQDSITPGEFIVGTTVEEGKTKINSWDNINEKAALIPSSAADYDETLNNNFQFN
 50 DNLQSYFSNFTQAVGKVKFVNKKAKQLTVDLPIDYYGQAEITIGITQYVTEQAEKYFDKLD
 EYEIRIKDGNTPRALISKTDDKEPQVHIYHN*

Sequence 2291

55 Contig_0735_pos_858_538,
 putative peptide of unknown function
 atgtttatgtgccacttctttaacagatctagcaatggattgaacttgacgttcattgtt
 cccgtgacagactacaaaataaaggggataaaacctcttggaatgatataagaatatta
 ggtcatcaagtgatctttgatgattgatttaataaccaagatcattatcaaatcgg
 galaaacttcaatttagcttgaattatgaagcactgtctcagagcatgtatatga
 aaat

ttaactaagttatatagtagtgattcaaaaatagaatcccttggttcagaacttcgatatg
cctatatattcccagtgctaa

Sequence 2292

5 MFMCHFFNRSSNGLNLTFFIVPVTDYKIKGIKPLGNDIRILGASSDHLMLDLNNQDHYQIG
DKLQFSLNYESQSMYMKNLTKLYSSDSKIESLVQNFDMPIYSQC*

Sequence 2293

Contig_0736_pos_4884_0,

10 is similar to (with p-value 3.0e-23)
>gp:gp|U93876|BSU93876_19 Bacillus subtilis aminoglycoside 6
-adenylyltransferase (aadK) gene, partial cds, and YrdA (yrd
A), YrdB (yrdB), hypothetical protein YrdC (yrdC), YrdD (yrd
D), hypothetical cytochrome P450 protein YrdE (yrdE), ribonu
15 clease inhibitor (yrdF), regulatory protein YrdG (yrdG), hyp
othetical protein YrdH (yrdH), hypothetical protein YrdI (yr
dI), amino acid transporter (yrdJ), YrdK (yrdK), LysR family
regulatory protein YrdL (yrdL), YrdN (yrdN), cation transpo
rt protein YrdO (yrdO), hypothetical protein YrdP (yrdP), Ly
20 sR family transcription regulator YrdQ (yrdQ), hypothetical
protein YrdR (yrdR) and hypothetical protein YrkA (yrkA) gen
es, complete cds. NID: g1934641. >gp:gp|Z99117|BSUB0014_140
Bacillus subtilis complete genome (section 14 of 21): from 2
599451 to 2812870. NID: g2634966.
25 gtgacaatcttagcgattgatattggagtggaatgtgggaatagcatcagcaattgtaaca
attgtgattatacttatttctgaagtgattcctaaatcaattgctgcaacatttctgat
aaaatttcaaaacttggtatcctatcattcatatattgttattgtactcaagccatt
acaatcttattaaacaagatgacagatgggtattaatcatttactatctcgaggccaacct
30 gttgaaaaaagattttctaaagaagaatttcgtacattattaaatattgcgggtagagaa
ggtgcatttaattgagatagaaaatactcgacttcaaaacgttatggactttgaacaattg
aaggttaaggatgttgataccacgcctcgtattaatgttgtagctttttcaaaggaaagta
acatatgacgaagcttatgatacagtgatgaataacccatatacaagatatccagiatat
gatgaaaatatagatgatcatcggcgtattccactcaaaatatttattagcttgaggt
35 aaaaaataagaggacgcaattactaattatgcatcaagccctttatttgtaaatgaacat
aatagggcagaatgggtattgcgtaaaatgaccgtttcacgaaaacatttagcgattgtt
ttagatgaatttgaggtacggatgctatcgatcgacgaagatttaatagaagagcta
cttggtatggatattgaggatgaaa

Sequence 2294

40 VTILADIGVNVGIIASAIVTIVIIILISEVIPKSIATFPDKISKLVYPIIHICVIVLKPI
TILLNKMTDGINHLLSRGPVEKRFSEKEIRTLNIIAGREGAFNEIENTRLQNVMDFEQL
KVKDVTTPRINVVAFSKEVITYDEAYDTVMNNPYTRYPVYDENIDDIIGVFHISKYLLAWS
KNKEDAITNYASSPLFVNEHNRAEWVLRKMTVSRKHLAIVLDEFGGTDIVSHEDLIEEL
LGMDIEDEX

Sequence 2295

Contig_0736_pos_3556_2282,

is similar to (with p-value 0.0e+00)
>sp:sp|P37949|LEPA_BACSU GTP-BINDING PROTEIN LEPA. >gp:gp|D8
50 4432|BACJH642_108 Bacillus subtilis DNA, 283 Kb region conta
ining main element. NID: g2627063. >gp:gp|X91655|BSLEPOFF_1
B. subtilis lepA and hemN genes. NID: g1122397. >gp:gp|Z99117
|BSUB0014_31 Bacillus subtilis complete genome (section 14 o
f 21): from 2599451 to 2812870. NID: g2634966.
55 atgtcaaaagatatattgcattaaagtaataaaaacgttactatatcaaagatacatagaa
gtgacaggttataaaagatgaaagcgagaaggataaaaatggataagcaagaacgatacaat
agaagagaaaatattagaaatttctccattattgctcatatagaccatggtaaatcgaca
ttagctgatcgaaattttagagaatacaaaaatcagttgaaactcgagaatgcaagatcaa
ttacttgactctatggatttggaagagaacgagggcatcactattaaactaaatgctgtt

cgattaaaatacgaagctaaagatggagaaaacttacacatttcatttgatagatacacca
 ggacatgtcgactttacatatgaggtttctcgctcattagctgcatgtgaaggtgcaatt
 cttgtagttgatgctgccaaggtatagaagcacaaccttagcaaacgtttatt'agca
 ttagataacgatttggaaacttttgccagttgttaataaaatagacttgccctgcagctgag
 5 cccgatagagtttaagcaagaattagaagatgttataggtatagatcaagaagatgtagta
 cttgcaagtgtcaagtcaaatataggtattgaagaaatttttagagaaaatagttgatgtt
 gtaccagcaccggagcgtgatccagaagccccacttaagcacttatctttgattcagaa
 tatgatccatacagaggagtaatatcttcaattcgaattattgatggtgtgtttaaagct
 ggagataggattaaaatgatggctaccggtaaagaatttgaagttacagaagtcggaatc
 10 aatacccttaagcaactaccggtagaagaattaacagttggtgatgtgggttatattatc
 gcaagtatcaaaaatgttgatgattctagagtaggtgacacaattacttttagctgaaaga
 cctgctgacaaaccgttacaaggatataaaaagatgaatccaatgggtattttgtggtcta
 ttccctattgacaataaagactataatgacctaaagagaagcttttagaaaaattacaact
 aatgacgcctccttagagtttgaaccagagtccttcacaagcacttggttttggatacaga
 15 actggatttttaggaatgttacatatggagattattcaagaaagaattgaaagagaattt
 ggtattgaactcattgcaacagcgcttcatcccgagaagtatcaatgcatgtgtgaac
 accattcgcttttaa

Sequence 2296

20 MSKIYCIKVIKTLLYQRYIEVTGYKDESEKDKMDKQERYNRRENIRNFSIIAHIDHGKST
 LADRILENTKSVETREMDDQLDSDMLERERGITIKLNAVRLKYEAKDGETYTFHLIDTP
 GHVDFTYEVSRLAACEGAILVVDAAQGIEAQTLANVYLALDNDLELLPVVNKIDLPAAE
 PDRVKQELEDDVIGIDQEDVVLASAKSNIGIEEILEKIVDVPAPDGDPEAPLKALIFDSE
 YDPYRGVISSIRIIDGVVVKAGDRIKMMATGKEFEVTEVGINTPKQLPVEELTVGDVGYYI
 25 ASIKNVDDSRVGDITLAERPADKPLQGYKMNPMVFCGLFPIDNKDYNLREALEKLQL
 NDASLEFEPESQALGFGYRTGFLGMLHMEIIQERIEREFGLIELIATAPSSRRSINACVN
 TIRF*

Sequence 2297

30 Contig_0737_pos_2511_2855,
 putative peptide of unknown function
 atgcgtagtgtaattagttggaaaatacatttatattcataatcacaaatcgccgtttat
 gttaaattagaattttctatatggctgtttttacttttactattagttccagatatattt
 atggttaggatatgtgattaatagaaaaacaggagttatgtttacaatttggacacacg
 35 tatatcacacctataattatcgcgctatttatatttatacattgatgaaaggttactatta
 cagattgcttttaatatggttagctcatattagtagtatggatagaacttttaggtttcggactc
 aatatcatcagatactgataaaacgataatacaaaaagatgtaa

Sequence 2298

40 MRSVIKLENTFIIITIAVYVKLEFSIWFLLLLLLVPDIFMLGYVINRKTGSYVYNIGHT
 YITPIIIALLYLIDERLLLQIALIWLAHISMDRTLGFLKYSSDSDKTIQKM*

Sequence 2299

45 Contig_0737_pos_6151_4991,
 putative peptide of unknown function
 gtgttttacgacacagacagtactgaagcgatgaaaagtcatatgagtgatttagtatta
 ggcaagcaagaacaaattgcttatatcaatcagttagaacgtggacttgaagaaaataaa
 attgaaagaaactctaattctaattagattaatcaagttgagaatgagcttggtcctgac
 gaaacctttgaaaaagaaaaggaatatacacacaagaagttttagaattacatgaaaaagag
 50 aacttgatgaaaagttaaaagaaacttttgaagaagaacaaacacaaaaaataaaaga
 caaaagtttttgagaataggatttatgttttgactattctatcagcagcactttctatt
 ttttcttttttactgcaaatcttatttttgggtataatatttgcctctattaactgtgatt
 tttgtagtaggtatcattttttctagatctaaagcagtagattatagcacagcaataagt
 caggaaattaatgatttagaaaaccaactcacgcaactgaaaaagaataaattcttgac
 55 ttcgatttagaatatcaacaagaagttcggaacaatggcgatgctaaaaaaaataaaa
 aaaatacttgaagaaaaacatcaatatatcaatcaatcattaacgactgcaaatgagcga
 ttagatagtttaaaacatagcatttttaactactaatagctggcataatagcaggactt
 ttatctggtgtagtcaaacaggttgggaggtaatgtcccaccccgacacactagtaga
 gatgccactaaccacactcaacaactcttacaattactaggaataccgtcaaatattacc

catctcacatacaatTTTTCTGAGCATGCATTACCTTGGATAAGTTTTATCGTACACTAT
 agTTTTCTATCGCTATTGCAATAATCTATATTTATATCGCAAAGAAATATACAAAAATC
 acactaggttatggtgcttttatttggtatagttatttggtattgtttttcatttaattctta
 atgccaattatgcatgctgaccgaatgcttttgatcaaccattttcagaacacctatca
 5 gaatTTTTTGACACATTGTTTGGATGATGGTTATAGAAATGGTCAGAAGGTATTTCTAT
 aatattcaattaaataataa

Sequence 2300

VFYDSTTEAMKSHMSDLVLGKQEQIAYINQLERGLEENKIERNSNSNEINQVENELVPD
 10 ETFEKKKEYTQQVLELHEKENLYEKLKETFEETQTKNKRQKFLRIGFIVLTILSAALSI
 FSFTANLIFGIIFALLTVIFVVGIIIFSRSKAVDYSTAISQEIENDLENQLTQLEKEYNLD
 FDLEYQQQVREQWRHAKKNKKILEEKHQYINQSLTTANERLDSLKHSIFKLLIAGIIAGL
 LSGVVKLGWVMPFPPTPSRDATNPQQLLQLLGIPSNITHLTYNFSEHALPWISFIVHY
 SE:ATA:YIYIYIAKKYTKITLGYGALFGIVIWIVFHLILMPIMHVVPNAFDQPFSEHLS
 15 EFFGHIVWMMVIEVMRRYFYNIQLNK*

Sequence 2301

Contig_0737_pos_3491_3075,
 putative peptide of unknown function

20 atggaattatagtagtgcctatgtttaataaatttttagttaacgacatagataaatca
 tcggaatggtatcaagagaatttaggttttaaaagtatttttaaatataaaatgaacaa
 aatcaaattttaattgagcattttacgattagcaaaatatcaagatttgatgttaatttct
 ggcaaacagtttggaagtcggtaatgcagttttatacaaatatacttgtaacaaatattcga
 attttaaaacaacgaataccttctcaatatatcggtgaagatcttgaagaaaaaccatgg
 25 aattctattgaaatgacaattaagatttagataatcatttaattacgcttacacaaagt
 aacataaaaaatgaagaatttaattgctttgatgcaacatacttcaaaaacattttaa

Sequence 2302

MELYSPMPFNKFLVNDIDKSSEWYQENLGFKSIFKFKNEQNQILMEHLRLAKYQDLMLIS
 30 GKQFEVGNNAVYTNILVNPINIRILKQRIQPSQYIVEDLEEKPWNSIEMTIKDLNHLITLTQS
 NIKNEEFNALMQHTSKTF*

Sequence 2303

Contig_0738_pos_2962_3540,
 is similar to (with p-value 4.0e-47)

35 >sp:sp|P42085|XPT_BACSU XANTHINE PHOSPHORIBOSYLTRANSFERASE (EC 2.4.2.-). >pir:pir|S51309|S51309 xanthine phosphoribosyltransferase - Bacillus subtilis >gp:gp|L77246|BACYACA_2 Bacillus subtilis (YAC10-9 clone) DNA region between the serA and
 40 kdg loci. NID: g1256615. >gp:gp|Z99115|BSUB0012_148 Bacillus subtilis complete genome (section 12 of 21): from 2195541 to 2409220. NID: g2634478. >gp:gp|X83878|BSXPTPBUX_1 B. subtilis xpt and pbuX genes. NID: g633168.
 gtggagtcgtaggacgaaaagtcaaagaagatggcgttgatcgatgagaaaattttg
 45 aaggtatagtggtatttttaaatcatcaaatgtagcaagttgatgaatgatgtaggtaaa
 acattttatgagtcctttcaagacgctggtattactaaaattttaactattgaagcttct
 ggtattgcgcctgctattatggcttcttttcattttgatgttccttgctatttgctaaa
 aaagctaaacctagctatttgaagatggcctttatagcacggatattcattcatttaca
 aaaaataaaacgagtagcagtcattgtatctgaagaatttttaggtgcagacgataaagta
 50 cttatcattgatgacttttagctaatggtgatgcttcgctaggtcttaattgacattgta
 aaacaagcaaatgacgacagctggcgtgggtattgtggtgaaaaagtttccaaat
 ggtcgccaacggttagaagatgcaggcttatatgtatcttcactttgtaaggtagcttca
 ttaaaaggcaataaggtaactcttttaggtgaagcgtaa

55 Sequence 2304

VESLGRKVKEDGVVIDEKILKVDGFLNHQIDAKLMNDVGKTFYESFKDAGITKILTIEAS
 GIAPAIMASFHDVPCFLFAKKAKPSTLKDGFYSTDIHSFTKNKTSVIVSEEFLGADDDKV
 LIIDDFLANGDASLGLNDIVQANATTVGIVVEKSFQNGRQRLEDAGLYVSSLCKVAS
 LKGNKVTLLEGA*

Sequence 2305

Contig_0738_pos_3579_4808,

is similar to (with p-value 8.0e-95)

- 5 >sp:sp|P42086|PBUX_BACSU XANTHINE PERMEASE. >pir:pir|S51310|S51310 xanthine permease - *Bacillus subtilis* >gp:gp|L77246|BACYACA_3 *Bacillus subtilis* (YAC10-9 clone) DNA region between the serA and kdg loci. NID: g1256615. >gp:gp|Z99115|BSUB0012_147 *Bacillus subtilis* complete genome (section 12 of 21):
- 10 from 2195541 to 2409220. NID: g2634478. >gp:gp|X83878|BSXPTPBUX_2 *B. subtilis* xpt and pbuX genes. NID: g633168.
- atgtatgcaggggctattcttcttctattattgtggggacaagcttaaaattttcagct
gaagaaattgcttcttagttactgttgatataatttatgtcggggtagcgacatttctt
caagcaataaaagtcacagggactggattaccgattgtactaggatgtacgtttactgcc
15 gttgcacctatgatactcatcggtcaaacgaaaggacttgatgttttatatggttcgctt
ttaatatccggtatcttagttgttttaattgcacctttttctcttatttagttaaatc
ttccacctgttgtaacaggaagtgtgtgacaattattggaatcaatttaagccagtt
gcaatgaattacttggcaggtggtgaaaggagcgaaaaactatggcgatactaagaattta
atattagggtggtgttacactactcattattcttattttgcaaagatttacaaagggttc
20 ttgaaatcaattgcgatacttataggattagcaaataggctgcttttagctggtatatt
ggaatgggtgatatacaacaagtgggtgatgcacattgggttgggttcctgtgccattc
agattttctggcttcggatttgatgtcagctcaatacttgatattttcattgttgagtt
gtaagtttaattgaatctactggtgtctatcatgcactgagtgaaattactggtagaaaa
ctagaaagaaaagattttcgaaaagggtacactgcggaagggtctagcaatcatttaggt
25 tcaatatttaatgcgttccttacactgcatattcccaaaatgtaggtcctgtttcttta
tcaggagctaaaaagaacaatgtgatataatggaatggttattctttactaatttgcggt
tgtatacctaaatttaggtgcttttagctaatattattccattgcccgggttttaggtggagca
atgatagcaatgtttggaatgggtatggcatacggcggttagtatatttggttaacattaat
ttccaaaatcaaaaataatttatttaattattgcaatttcagtaggggttaggtgctggtatt
30 agtgcagtacctcaagcatttaaggatttaggagaacaatttgcttggttaactcaaaat
ggtatagtgttggcgcaatttctgcaatcatcttaaaattcttttttaattggtataaag
tataaacaactgaagaaaatgtgaaataa

Sequence 2306

- 35 MYAGAILVPIIVGTSLKFSAEIAYLVTVDIFMCGVATFLQANKVTGTGLPIVLGCTFTA
VAPMILIGQTKGLDVLVYGSLLISGILVVLIAFFFSYLVKFFPPVVTGSVVTIIGINLMPV
AMNYLAGGEGAKNYGDTKNLILGGVTLIIILILQRFETKGLKSIAILIGLAIGTALAGIF
GMVDIKQVGDAHWFGFPVFRFSGFGFDVSSILVFFIVAVVSLIESTGVYHALSEITGRK
LERKDFRKGTYAEGLAAILGSIFNAFPYTAYSQNVGLVSLSGAKNNVIYGMVILLICG
40 CIPKLGALANIIPLVPLGGAMIAMFGVMAYGVSIILGNINFQNNLLIIAISVGLGAGI
SAVPQAFKGLGEQFAWLTONGIVLGAISAILNFFFNIGIKYKQTEENVK*

Sequence 2307

Contig_0738_pos_4846_6312,

is similar to (with p-value 0.0e+00)

- 45 >sp:sp|P21879|IMDH_BACSU INOSINE-5'-MONOPHOSPHATE DEHYDROGENASE (EC 1.1.1.205) (IMP DEHYDROGENASE) (IMPDH) (IMPD). >pir:pir|S12623|DEBSMP IMP dehydrogenase (EC 1.1.1.205) - *Bacillus subtilis* >gp:gp|X55669|BSIMPDE_1 *Bacillus subtilis* guaB gene for IMP dehydrogenase. NID: g39958.
- atgtgggaaaataaatttgctaaagaatctttaacattcgacgacgtgttactcattcca
gctgcatcagatgttttaccaagcgatgttgacttaagtgtcaaattatcagataagatc
aagttaaacattcctgttatctcagcaggtatggatacagtaactgaatcaaaaatggca
attgctatggctcgacaaggcggttaggtgttattcataagaatatggcgctcgaagag
55 caagctgatgaggtacaaaagggttaaacgttcagaaaatggtgttatttctaaccggttc
ttcttaacaccgggaagaaagtgtgtatgaggctgaagcattaatgggtaaaataccgtatc
tctggtgtaccattgtcgataatcaagaggatcgcaagttgattgggattttaacaaat
cgtgatttacgtttattgaagattttcaattaaaatcagatgtaatgacgaaagat
aatttaataacagctccagttggtacgacttttagatgaagccgaggctattcttcaaaaa

cataagattgagaaacttccattagtagaaaatggctggttagaaggattaatcactatt
 aaagatattgaaaaagtacttgaattcccatatgcagctaaagatgaacatggcagattg
 ttagctgcggcagcaatcggtacgtctaaagatactgaaattcggtgcacaaaaactagtt
 gaagctggcgtagatgcattaattattgatacagctcatggctcattctaaaggcggtatt
 5 aatcaagttaaacacatcaaggaaacatatcctgaaattactgttgctcgctggtaacgta
 gcgactgcagaggcaacacgtgctttatttgaagcgggtgccgatggtgttaaagtaggt
 attggtccaggctcaatttgcacaacacgtgttggtgcaggtgtaggtgtgcctcaaatt
 acagcagtttatgattgtgtacagaagcccgtgaagcatggtgaaggctattattgctgat
 ggtggtattaagttctcaggtgatattatcaaagcattagctgctggtggtcatgcggtt
 10 atgttaggtagttttagctggtacagaagaaagtcctggtgcaactgaagtattccaa
 ggtagacaatataaagtttatcgcgccatgggatcttttaggtgctatggaaaaaggttca
 aatgatcggttacttccaagaagataaaaacaccaagaaaaattgttcctgaaggtattgaa
 ggtcgtacagcttataaaggaccattacaagatacaatttatcaacttatgggtggcggt
 agagctggcatgggttatactgggttcagaaaacctaaaaaaattacgtgaagaagcacia
 15 ttacacgtatgggaccagctggcttagctgaaagtcacctcataatgttcaaattacg
 aaagaatcaccaaactattctttctag

Sequence 2308

MWENKFAKESLTFFDDVLLI PAASDVLPSDVL SVKLSDKIKLNIPVISAGMDTVTESKMA
 20 IAMARQGGLGVIHKNMGVEEQADEVQKVRSENGVISNPFLLTPEESVYEAALMGKYRI
 SGVPIVDNQEDRKILITNRDLRFIEDFSIKISDVMTKDNLITAPVGTTLDEAEAILQK
 HKIEKLPVLENGRLEGLITIKDIEKVLEFPYAAKDEHGRLLAAAAIGTSKDTAIRAQKLV
 EAGVDALIIDTAHGHSKGVINQVKHIKETYPEITVVAGNVATAEATRALFEAGADVVKVG
 IGPGSICTTRVVAGVGPQITAVYDCATEARKHGKAIADGGIKFSGDIKALAAGGHAV
 25 MLGSLLAGTEESPGATEVFQGRQYKVYRGMGSLGAMEKGSNDRYFQEDKTPRKFPVPEGIE
 GRTAYKGPLQDTIYQLMGGVRAGMGYTGSENLKKLREEAQFTRMGPAGLAESHPHNVQIT
 KESPNYSF*

Sequence 2309

30 Contig_0738_pos_6473_0,
 is similar to (with p-value 5.0e-94)
 >gp:gp|U51115|BSU51115_8 Bacillus subtilis CotA (cotA), GabP
 (gabP), YeaB (yeaB), YeaC (yeaC), YebA (yebA), GMP syntheta
 se (guaA) genes, complete cds, and AIR carboxylase I (purE)
 35 gene, partial cds. NID: g2239287. >gp:gp|Z99107|BSUB0004_83
 Bacillus subtilis complete genome (section 4 of 21): from 60
 0701 to 813890. NID: g2632866.
 atgactatggaaatggcgaaagagcaagagctgattcttgttttagactttggtagccaa
 tataaccagtttaattacgcgtcgtatccgtgagatggcggtttatagtgaattacatgat
 40 cagcaaatcttctattgaagaaattaaacgaatgaatcctaaaggatcattctttcaggt
 ggtccaaattcagtgatgaagaggggtcatttaccatcgaccctgaaatttacaattta
 ggtattccagtttttaggtatatgttatggtatgcaattaacgactaagcttttaggtggt
 aaagttgagcgtgccaatgagcgtgaatatggcaaagctacaattaacgctaaatcagat
 gaacttttcttttgcttaccttctgaacaaacagtatggatgagtcattctgataaagta
 45 attgaaattcctgaaggatttgaagtgattgcagatagtcgaagtactaattatgcagct
 attgaagataaaaaacgtcgatttacggtgtacaattccatccagaagtacgtcacact
 gaatatgggaacgacttactaagaaaacttcgtccgctggtttgtaattgtacaggtgaa
 tggacgatggagaatttcattgaaattgaaattgagaaaatccgtcaacaagtaggtaat
 cgtaaagattatgtgcaatgagtggtggagtagattcatccgtagttgctgtactttta
 50 cataaggcaatcgcgaccaattaacatg

Sequence 2310

MTMEMAKEQELILVLDGFSQYNQLITRRIREMGVYSELHDHEISIEEIKRMNPKGIILSG
 GPNSVYEEGSFTIDPEIYNLGI PVLGICYGMQLTTKLLGGKVERANEREYKATINAKSD
 55 ELFFGLPSEQTVWMHSDKVIIEPEGFEVIADSPSTNYAAIEDKKRRIYGVQFHPVVRHT
 EYGVNLLPNEFVRRVCNCTGEWTMENFIEIEIEKIRQQVGNRKVLCAMSGGVDSSVVAVLL
 HKAIGDQLTX

Sequence 2311

Contig_0738_pos_2368_1961,
putative peptide of unknown function
atggcaaaaattactgtagtgaataaccaagatgaattatataaaagtcataatcaaaaa
aaatctgaagggtatttagagacagaattagctgtatcagtaaaagtaagttgcactta
5 gatgatttacacaactctcaaactctcgtaaatggctacaagtggctcatttagtgaccgt
atgtctcgttacttacaggtgaagatggagaagaacagatattatctcgttatgattta
actgacaatgaactagaaggatataaacaagatattttaacgataaaaatgctcggtgtt
gcaaacagtgaccgttcttctcatgatgaagttgaagataataatgctgcatataaagaa
gtggataactactcattatgccgcagagtcctgaagggcctaaagcataa

Sequence 2312
MAKITVNNQDELYKVINQKKSEGYLETELAVISKSKLHLLDDLHNSQISLMATSGSFSDR
MSRLLTGEDGEETVLSRYDLTDNELEGYKQDILNDKMLVVANSRSSHDEVEDNNAAYKE
VDITHYAAESEGPKA*

Sequence 2313
Contig_0738_pos_1833_1282,
putative peptide of unknown function
atggaatttaaagtgatagaatcagctaaagatccattatttaacgaggcactcaaatta
20 tatgatgataaattggatattggttagatgaagatagtaaaatttttaaacgctcactt
gaaaataataaaacagaaaaatgattacgcctttatcggttggaattgaaaaccagactgta
gttagcttagcaactgcacattacgaagcaacaaccaattctgcatttttaatttactta
atcgcaaaagaaagccccaatcatgatgaaagaatgtccttaacttttagaggcaatagaa
aaacaattaaacctttttatcacaagaagttcataatagagatattaatttcatcatgtta
25 gaagttccaaaagaaccctcgactgctaacatcgatgacaagctcagaaatgcactagaa
catcgctcgtaatttctctttgaaaatcaatttgaaaagcaggacgacattgactatatt
catccaaaccaaaccaaaaggaaacgcccaaaaagtagatttatttattaaagcgaac
attgcatatag

Sequence 2314
MEFKVIESAKDPLFNEALKLYDDKLDIGLDEDSKIFKRSLENNKTENDYAFIVGIENQTV
VSLATAHYEATTNSAFLIYLIAKESPNHDERMSLTLEAIEKQLNLLSQEVHNRDINFIML
EVPKEPSTANIDDKLRNALEHRRQFLFENQFEKQDDIDYIHPNQNKETPQKVDFIKAN
IEL*

Sequence 2315
Contig_0741_pos_1807_2565,
putative peptide of unknown function
atggatgacttgaaacaaaaatcaatcttctaacgaaaaacctaaggaataaaaaataatt
40 aatattttgatattcatcggaatgatttttatttaattcaaatacctattggcgtgtcacta
atagctttacctttttcagtgaaattcagtaagttaacatccatcgcatthaagtatgcta
ataactgggtacagcactattaatcatatgggttagttaggaattattattttgagtcataca
tatgaaagacaatatcaatcaatgaggggaaaagatatctttattaatattgggtttctg
gtattatcaatgggttttagtattctaagtagtgattaatggatcatatttactggcaac
45 gatactacagcgaatgagaaagaaatcaatgaaagtttagatttacttttcaaaaaagac
catttaccacatatattcaattgttgcaactgttggttttaatgatatgtattataggtccg
taatttagaggaattactcttccgaggaatttttaagaaacattatttatgaaatatcga
ttttggcaccattcattatatcttctattatttttagttcacaacatttatcaa.aaat
atattttcatatgcaatttttttctaattgggttggtattataccttgccataaacaga
50 agacgtaatatcaaagatagtagtggttcacatggtgaataattctgtttcaacatta
ccggtatttgggtggttatttatggctatatttttagatag

Sequence 2316
MDDLKQNOSSNEKPKGNKIINILIFIGMILLIQIPIGVSLIALPFSVKFSKLTSLALSML
55 ITGTALLIIWLVNRNYLSHTYERQYQSMRGKDIFINIGFLVLSMVFSILSSVLMVIFTGN
DTTANEKEINESLDLLLQKDHLPHISIVATVVLMICIIGPYLEELLFRGIFKETLFMKYR
FWLPFISSII FSSQHLSTNIFSYAIYFLMGCVLAYNRRRNKIDSMVMHMLNNSVSTL
PVFVGYLWLYFR*

Sequence 2317

Contig_0741_pos_3896_4231,

putative peptide of unknown function

gtgtatatcatgtatgagaatatacaaaatattgaagattttaccgatttataaatgca
 5 caccgttttagctgttgttcatattatgagagataattgtacagtatgtcatgcagtatta
 ccccaaattcaagatttactaaaggactatccgaaagcacaattaggtgtgattaatcaa
 tctaattgtgaagctattgccggagaactttctatttttacagtacctgttgatttaatt
 tttttgaaagggaaagaaatgcatagacaagcacgtttttatcgatatgcaatcgtttgaa
 10 aaacaattgtatataatgcaaaatgccatcgattaa

Sequence 2318

VYIMYENIQNIEDFYRFINAHALAVVHIMRDNCTVCHAVLPQIQDLLKDYPKAQLGVINQ
 SNVEAIAGELSIFTVPVDLIFLKGKEMHRQARFIDMQSFEKQLYIMQNAID*

Sequence 2319

Contig_0741_pos_4399_5262,

is similar to (with p-value 1.0e-31)

>sp:sp|P42978|YPJC_BACSU HYPOTHETICAL 23.6 KD PROTEIN IN QCR
 C-DAPB INTERGENIC REGION. >gp:gp|L38424|BACJOJC_1 Bacillus s
 20 ubtilis dihydronicotinamide reductase (jojE) gene, complete cd
 s; poly(A) polymerase (jojI) gene, complete cds; biotin acet
 yl-CoA-carboxylase ligase (birA) gene, complete cds; jojC, j
 ojD, jojF, jojG, jojH genes, complete cds's. NID: g755600. >
 gp:gp|L47709|BACYPIA_9 Bacillus subtilis (clone YAC15-6B) yp
 25 iABF genes, qcrABC genes, ypjABCDEFGH genes, birA gene, pan
 BCD genes, dinG gene, ypmB gene, aspB gene, asnS gene, dnaD
 gene, nth gene and ypoC gene, complete cds's. NID: g1146223.
 >gp:gp|Z99115|BSUB0012_191 Bacillus subtilis complete genom
 e (section 12 of 21): from 2195541 to 2409220. NID: g2634478

atgaaggaggccaaaatgaagaagcaacagaataaagtccatatgattaatatatcttta
 tgccttataggacattacttattgcaatcgagtaaatagttttgttataccgggcaat
 ttaggtaggggtgttctataggcctttcatttaattctgaattatactctaggtatttca
 30 cccgcactcagttccttcatcattaacgcaatattaattattgtaggttgaaatttctc
 agtagaacgacagcgatttacactgctataacaatcaccgcaagttcaatttttcttgat
 ttaactcacacattcggattaggtattcatgataattttattaattcaatttttgcaggt
 ttaattgttggaatcggttctgggttagtaattactgctcatagtagcattgggtggtaca
 35 tctgtcattgcacgtatcatttcaaaatagtgagatgaagacgtcacaaagcactactt
 atattagatgcaatcatcgtgttatcatttatcgttgtttacctataacgaatgatta
 40 tatactattgtgctatttattgttgaaaaatcaatgtcttttgttgaaggattt
 aatcctaaaaagctgtgacagttatttcaaaatataataaagaaatcagtgctgatata
 tatgaaatgactggaagaggggcaaccttattaagtggtaaagggtgcttaccaaaaaagt
 gatacagaagttctatatgccgtggtatcccaaatcaagttggagcaataaaaaagatt
 45 gttaataatgatgaaaatgccttttttagtgattcatgatgtgcgtgatgtcttaggt
 aatggatttattaatattaataa

Sequence 2320

MKEAKMKKQONKVHMINISLCLIGTLLIAIAVNSFVPGNLGEGGSIGLSLILNYTLGIS
 PALSSFIINAILIIVGWKFLSRRTTAITYTAITITASSIFLDLTHTFGLGIHDFINSIFAG
 50 LMLGIGSGLVITAHSTLGGTSVIARIISKYSEMKTSQLILDAIIVLSFIVVLPITNVL
 YTIVMLFIVEKSMSEFVVEGFNPKKAVTVISKYNKEISADIYEMTGRGATLLSGKGAYQKS
 DTEVLYAVVSQNVGAIKKIVNQYDENAFLVIHDVRLVGNLFINIK*

Sequence 2321

55 Contig_0741_pos_7460_8320,

is similar to (with p-value 8.0e-77)

>gp:gp|Z92954|BSZ92954_1 B.subtilis yws(A,B,C,D,E,F,G) and g
 erBC genes. NID: g1894764.

gtgggttaacaaaaagcaaatgccctagttaatgagtgatggctgtaaatccagattat

caaattacctttcaaaatgatttagtaaaagcaaataatcgggtgaattgttaatgtgatg
 gaagaccatatggatgtcttaggaccgacacttaagatgtagcgcaagcttttactgca
 acaattccatataacgggaaattagttgtaatgaaagataactatactagtttctttgca
 aaggaagctaaaaagcgtaattcagaactcattgtttagataaagacgtcataaccagaa
 5. tcatatttacggaagttcgattattttagtatttcctgataatgtagctattgtgttagga
 atagcgcaagcagttggtgtagatgaagaaactgcattacaaggatgttaaatgcacca
 gccgatccaggtgctgttagaattaaatatttccatgcaaatcgcaaaaaaatgtattt
 gttaatgcattcgctgctaatagaaccgcagtcctacaaaagcgattttaaataaagtggaa
 tcatataaattatccatacgataagaaaaataatcattctcaattgtcgttcagatagggt
 10 gatagaacacaactctttgttgataactttttaggtgaagtcgattacgatgttctcatt
 tgtacaggaaaaagtcacaaaatggtgacacagtttatggaaactatgccagaaaaaaca
 tatatcaattatgaaggacgagactttttagagattgaaaaaggatttctacatgaagct
 gagaatgcacttgattttgtgtaggaaacatccacggcccggtggtagaatagcggaa
 ttcatagaagggatagaataa

15

Sequence 2322

VVKQKANALVNECMVNPYQITFQNDLVKANIGVIVNVMEDHMDVLGPTLKDVAQAFTA
 TIPYNGKLVVMKDNYSFFAKEAKKRNSLIVVDKDVIPESYLRFKFDYLVFPDNVAIVLG
 IAQAVGVDEETALQGMLNAPADPGAVRIKYFHANRTKNVFVNAFAANEPOSTKAILNKVE
 20 SYNYPYDKKIIILNCRSDRVDRTQLFVDNFLGEVDYDVLICTGKSTQMVTQFMETMPEKT
 YINYEGRDFVEIEKGILHEAENALVFCVGNHGPGRGRIAEFIEGIE*

Sequence 2323

Contig_0741_pos_8322_8774,
 25 is similar to (with p-value 2.0e-43)
 >gp:gp|Z99122|BSUB0019_86 Bacillus subtilis complete genome
 (section 19 of 21): from 3597091 to 3809700. NID: g2636029.
 >gp:gp|Z92954|BSZ92954_2 B.subtilis yws[A,B,C,D,E,F,G] and g
 erBC genes. NID: ql894764.
 30 atgatagggttcagaattatatttctccttatttcgtaggtgtcgtactcagtttgatattt
 gctgagaaatttgggattaatccagcagggttagtcgttccagggtatttagctttgatt
 tttgatcaaccgatcatgttgttatcagttatattcatttagttgcttaactattttatc
 gtaagcaacggtattagtaagtggttattttatatggtagaagaaaaattcgctgccatg
 atactgacgggaatggtgattaaatttatatttgatctcttgtaaccattgacccattt
 35 gaaatggttgaagtttcagggtatagggtgtgtcattcctggtattattgcaatacaatt
 caaaaacaagggtgtagtcattacactttctacaacaatggtattaacatgtattacatat
 atcatcttatttttatatagttttattaattaa

Sequence 2324

40 MIGSELYFSLFVGVVLSLIFAEKFGINPAGLVVPGYLALIFDQPIMLLSVLIISCLTYFI
 VSNIGISKWVILYGRKFAAMILTGMVIFKIFDILLYPLTFEMVEVSGIGVVIPGIIANTI
 QKQGVVITLSTTMLLTCITYIILFLYSFIN*

Sequence 2325

45 Contig_0741_pos_8798_9400,
 putative peptide of unknown function
 atgacaaaagaaaaaacgtttatcgccctagtgagtggttgcctaaacaatctaaaagacat
 aaaaggaaaaatacactttacacggcaattgtacttttagtagcggttagttctactcata
 tttgctgttaaatcaatacagaagtagaacctgtaaaaagtatacagagagacaaagatagc
 50 attcgtatcacctatttaggtaacgtcactttaaataaacatattcgacaaactaaattg
 aatgatgttttttaagggtattcaagatacttttagatcatagtgatttttcaacagytcca
 ttaatagtaaatgatttttcaagaaatcaaaaagataacataaataaaaaatattgaaat
 atcatgtttctacgcaagcataatgttaaaagtgttaacttaataacgaatctatggat
 aatattcaagcgacagcaatgatgagaaaaatagattcccaagcagggtataattttta
 55 acaggtaatggttcaaatccaattaatagtaaaactgtacaacaagacattaaaggtaaa
 aaaatagctaaaacataccatattgttgcagactatctaattgatgtaaatccggtattgt
 taa

Sequence 2326

MTKKRLSPSEWLLKQSKRHKRKNLTLYTAIVLLVALVLLIFAVKSIQVEPVKSDTRDKDS
IRITYLGNHTLNKHIRQTNLNDVFKGIQDTLDHSDFTGSLIVNDFSRNQKDNINKNIEN
IMFLRKHNKSVNLINESMDNIQATAMMRKIDSQAGYNFLTGNNGSNPINSKTVQQDIKKG
KIAKTYHIVADYLIDVNPDC*

5

Sequence 2327

Contig_0741_pos_10400_10837,

putative peptide of unknown function

atgagcattgaatttagacatcaatcatgggttgacaatcagtataaagaacaaacttta
10 tctttcttaacacaacatcaaattcattcatgcagtggttagatgaacctcaagttaaagag
gggagcgttccttttagtaaataggattactagtgaattgctttgtacgttatcatgga
cgtaatcattatggttggaactaaaaagatatgactgatcaagaatggcgagatgtaaga
tatttatagattatagcgatgatgagttagctgacttggctcgtaaaagtcgaaatactt
aatcaaaaggctaagaaagtatatgtaatttttaataaactctggcggtcatgcagct
15 aataatgctaaaaagtatcaaaatatttttagacattgattatgaaggttagcaccgcaa
caattaaaactatttttaa

Sequence 2328

MSIEFRHQSWFDNQYKEQTLSTLQHQIHAVVDEPQVKEGSVPLVNRTSEIAFVRYHG
20 RNHYGWTKKDMTDQEWDRVRYLYDSDELADLARKVEILNQKAKKVYVIFNNNSGGHAA
NNAKKYQNILDIDYEG LAPQQLKF*

Sequence 2329

Contig_0741_pos_11189_0,

25 putative peptide of unknown function

gtgtttatgatattgtgtcgcgatattactgatgattcgtcacaaaatcaaacctttttaa
atttttgacaaaacctaaatatgcgcgtacatatgttgatgctgaagggaacataccgt
tatagtgtaccacccctgtttgcttttataacaacgttatttattgggctattaacagga
ctgtttggcataggtggaggtgcattgatgacccctcttatgctcatcgtcttttagattt
30 ccaccacatggtgcagtaggcacaaagtatgatgatgattttcttttcaagtgtgatgagt
tcaatagggcacatctttcaaggacatgtggcttgggctattctatcattctcattatt
tcaagtgttataggtgcacaaattgaggacttcatttctccacatcaggattgaaatat
gagcgcacatcaatataacttatcggccaacttcatagaagaagaatattattatgatgac
tcttctt

35

Sequence 2330

VFMIFVSILLMRHKIKPFKIFDKPKYARTYVDAEGKTYRYSVPPLFAFITTLFIGLLTG
LFGIGGGALMTPLMLIVFRFPPHVAVGTSMMMIFFSSVMSSIGHIFQGHVAVGYSIILII
SSVIGAQIEDFISPTSGLKYERHQYNLYGQLHRKEYYYDDSSX

40

Sequence 2331

Contig_0741_pos_9875_9345,

is similar to (with p-value 1.0e-58)

>sp:sp|P23920|SYM_BACST METHIONYL-TRNA SYNTHETASE (EC 6.1.1.
45 10) (METHIONINE--TRNA LIGASE) (METRS). >pir:pir|S16682|S1668
2 methionine--trna ligase (EC 6.1.1.10) - Bacillus stearothe
rmophilus >gp:gp|X57925|BSMETSG_1 B.stearothermophilus metS
gene for methionyl-trna synthetase. NID: g39988.

gtggtaagagattacttaatgcgtgaattaccgtttggtctgatggcgtatttacaccg
50 gaagcctttgttgaagaacaaattacgatcttgcaatgatttaggtaatttagtgaat
cgtactatctctatgataacaaatatttccacggcgaattacctgcataccaaggtcca
aaacatgaattggatgaaaaaatggaagcgatggcgcttgaaactgttaaatcattcaat
gataatatggaagtttacaattttctgttgctttatcaacagtatggaattttattagt
cgtacaaacaaatatattgatgaaactcaaccttgggttcttgcaaaagatgaaatcaa
55 cgtgagatgcttggttaattgtaatggcacatcttgtcgagaacattcgtttcgctacaatc
ttattacaaccattcttgacgcgatgcacctagagagatatattaagcaacttaattaac
aatccggatttacatcaattagatagctctgcaacaatatggtatgtttttag

Sequence 2332

VVRDYLMLRELPGSDGVFTPEAFVERTNYDLANDLGNLVNRTISMINKYFHGELPAYQGP
 KHELDEKMEAMALETVKSFNDNMESLQFSVALSTVWKFISRTNKYIDETQPWVLAKDENQ
 REMLGNVMAHLVENIRFATILLQPFLLTHAPREIFKQLNINNPDLHLQDLSLQQYGMF*

5 Sequence 2333
 Contig_0741_pos_6714_6397,
 putative peptide of unknown function
 atggagaagtcagttaaacttgctgttgcatctatctagcaattattttaattatttgt
 agtattttacctcgcatttatacttattggaagcttaaatggtaaagacatgagtaattct
 10 gttttagatactgatcactctcgtatcaacaatacttcaagaaacagtaacgaagatgtt
 acgtcatcaaataatgagtcacaacaatacaaaagcgactcatttgcaaaactctgaatat
 aaagctattaacataaacgaagcatttaaaaaataataagcaaattaaaaaagcgaattcg
 agttatcaatactattga

15 Sequence 2334
 MEKSVKLAvgiylailiicsiylafiligslngkdmsnsvldtdhsrinntsrnsnedv
 TSSNNESNNTKAHSFANSEYKAININEAFKNNKQIKKANSSYQYY*

Sequence 2335
 20 Contig_0741_pos_2363_2061,
 putative peptide of unknown function
 atgaatggtagccaaaatcgatatttcataaataatgtttctttaaaaattcctcggaag
 agtaattcctctaaatacggacctataatacatatcattaaaaacaacagttgcaacaatt
 gaaatatgtggttaaaggtctttttgtaaaagtaaatctaaactttcattgatttctttc
 25 tcattcgctgtagtatcggttgccagtaaatatgaccattaatacactacttagaatacta
 aaaaccattgataataaccagaaaaccaatattaataaagatatcttttccctcattgat
 tga

Sequence 2336
 30 MNGSQNRYFINNVSLKIPRKSNSKYGPPIIHIKTTVATIEICGKWSFCKSKSKLSLISF
 SFAVVSPLPVNMTINTLLRILKTDNTRKPILIKISFPLID*

Sequence 2337
 Contig_0743_pos_1628_1969,
 35 putative peptide of unknown function
 atgttcacaatcaaacctgtagaaatagctacaactaaaatggcgataacaaatgagact
 acaaattttttcttgaatagtttcgataataatacaacttctggaatacttgaccagca
 ccaccaatgatcaaagcaacaactgtcccaagcgacattccttttgaaagtaatgcttca
 gctataggttaacattgtttcaggtctaataacattggaatgcctataacagatgcaatg
 40 aatacagatataacgccatcaccacttgctgtattttgtaataaatgtttcgggtacaaag
 ccataatgaacgctccaataaacacaccaataaataggttaa

Sequence 2338
 45 MFTIKPVEIATTKMAITNETTNFFLNSFDNNTTSGILAPAPPMIKATTVPSDIPFEANAS
 AIGNIVSGLIYIGMPITDAMNTDITPSPLAYFVINVSGTKPYMNAPINTPINR*

Sequence 2339
 Contig_0743_pos_3389_3778,
 is similar to (with p-value 2.0e-16)
 50 >gp:gp|AF016485|AF016485_124 Halobacterium sp. NRC-1 plasmid
 pNRC100, complete plasmid sequence. NID: g2822278.
 atgttaattctataatagcaaattaattaatgaggtgagtataatgttaaatattgaaata
 tacggaagagctatgtgctgttctactggagtatgtggtccagaaccgatgaaaacta
 ataaaagcgaaccaaataatgaatatttaaagcaaaatcaaataagaagttcaacgttat
 55 aatatgaataataatccgaatgaattcattaaaaatcaagaagttattcgtttaattcaa
 gaaaaaggtgatgaagttttaccaatcacttttatagaaggcggtatagctaaaacgggc
 gcttatattaccaagaagaagccgatgaaattattacagttaatcaaatgagaaatgga
 ggatgctgtggtggagatggatgctgttaa

Sequence 2340

MLIYNSKLINEVSIMLNIEIYEEAMCCSTGVCGPEPDETLIKANQINEYLKQNIQIEVQRY
NMNNPNPEFIKNQEVIRLIQEKGDVLPITFIEGGIAKTGAYITQEEADEIITVNQMRNG
GCCGGDGCC*

5

Sequence 2341

Contig_0743_pos_4248_4676,

putative peptide of unknown function

atgcttgaattaccttctgcatggacagattatTTAAATACAACGAGTAATGACGCTTCT
10 tgccttaggtcaattatcagggtttaaataGAAAATAGAGTTAAATATAATTcagcacttgaa
aaactacgtaaccaagatgatacgaccatgatgttagttgcgagacactcactcttct
atatatgaaattcaaagagcgcaacaagaattcatcatccactggtggtgtgactacttc
gcctgtatcaggatttttaactcctggtttacctggaacgtcctcttggtacctttcgg
tgcatttggatcaaattcactcttatggcctggcttgatttcttcgccaccataatgaac
15 gatttcactcactggttgtttgttactttttctgttggttcaccttcgccaactttttc
ccctgttaa

Sequence 2342

20 MLELPSAWTDYLNNTSNDASCLGQLSGLNENRVKYNSALEKLRNQDDTTMMLVARPTHSS
IYEIQRAQQEFIIHWWCDYFACIRIFNSWFTWNVLLATFRCIWIKFILMAWLDFFATIMN
DFIYWLFCYFFCWFTFANFFPC*

Sequence 2343

Contig_0743_pos_5242_4406,

25 putative peptide of unknown function

gtggatgatgtgacaaaatatggtccagttgatggagatccgattacgtcaacggaagaa
attccgtttgataaaaaacgcgaatttgatccaaacttagcgccaggtagacagagaaagtc
gttcaaaaaggtgaaccagggaacaaaaacaattacaacaccaacaactaagaaccatta
acaggggaaaaagttggcgaaggtgaaccaacagaaaaagtaacaaaaacaaccagtggt
30 gaaatcgttcattatggtggcgaagaaatcaagccaggccataaggatgaatttgatcca
aatgcaccgaaaggttagccaagaggacgttccaggtaaaccaggagttaaaaaccctgat
acagggcgaagtagtcacaccaccagtggtgatgtgacaaaatatggtccagttgatgga
gatccgatcacgtcaacggaagaaattccgtttgataaaaaacgcgaatttgatccaaac
ttagcgccaggtagacagagaaagtcgttcaaaaaggtgaaccagggaacaaaaacaattaca
35 acaccaacaactaagaaccattaacaggggaaaaagttggcgaaggtgaaccaacagaa
aaagtaacaaacaaccagtagatgaaatcgttcattatggtggcgaagaaatcaagcca
ggccataaggatgaatttgatccaaatgcaccgaaaggttagccaagaggacgttccaggt
aaaccaggagttaaaaatcctgatacagggcgaagtagtcacaccaccagtggtgatgaa
ttcttgttgctctttgaatttcatatatagaagagttagtaggtctcgcaactaa
40

Sequence 2344

VDDVTKYGPVDGDPITSTEEIPFDKKREFDPNLAPGTEKVVQKGEPTKTITPTTKNPL
TGEKVGEGETPEKVTQKQPVDEIVHYGGEEIKPGHKDEFDPNAPKGSQEDVPGKPGVKNPD
TGEVVTTPVDDVTKYGPVDGDPITSTEEIPFDKKREFDPNLAPGTEKVVQKGEPTKTIT
45 TPTTKNPLTGEKVGEGETPEKVTQKQPVDEIVHYGGEEIKPGHKDEFDPNAPKGSQEDVPG
KPGVKNPDVTGEVVTTPVDDEFLLRSLNFIYRRVSRSRN*

Sequence 2345

Contig_0743_pos_2507_1623,

50 putative peptide of unknown function

atgttagattcaattatggaatttataaaaaacatttgatggtgttttttgaattatta
atactattcatcatcgtagctttattgtaagtataatccagcagatagtttcagaagaa
aaaataaaacactttttaagtaaacctaataagctattaattataattttgggaatggct
tttgggtgcgatgacaccattttgttcttcttacaattcctatacttgacaggtttatta
55 aattctaaagttccatttggccctgcaatgagtttttaattgcgtcacctttaatgaat
ccattaatgatatttatgttatggccttattaggttggaagttgctgtgtgttacttt
attttactagcactcttttagtctttaaagggtctagtttttcaaaaatgaatttagct
gaaacttataaaggagtaaatgttaaaggcgatggattttttgctaataaaatgggatct
cgttttaacaagcattaaatgatgcgtgggcatttttatatccaatgcttccttaccta

5 tttattggtgtgtttattggagcggttcatatatggctttgtacccgaaacattttattaca
 aaatacgcaagtggatggcggttatatctgtattcattgcatctgttataggcattcca
 atgtatatttagacctgaaacaatgttacctatagctgaagcatttagcttcaaaaggaatg
 tcgcttgggacagttgttgccttgatcattggtggtgctggtgcaagtattccagaagtt
 gtattattatcgaaactattcaagaaaaattttagtctcatttgttatcgccatttta
 gtttagactatttctacaggtttgattgtgaacatcggtatttta

Sequence 2346

10 MLDSIMEFIKTFVMLFFELLILFIIVSFIVSIIQQIVSEEKIKHFLSKPNQAINYILGMA
 FGAMTPFCSCSTIPILAGLLNSKVFPFGPAMSFLIASPLMNPLMIFMLWALLGWKVAVVYF
 ILLALFSVLTLGLVFSKMNLAETYKGVNVKGDGFFANKMGSRFKQALNDAWAFLYPMLPYL
 FIGVFIGAFIYGFPETFTKYASGDGVISVFIVSVIGIPMYIRPETMLPIAEALASKGM
 SLGTVVALIIGGAGASIPVVLLSKLFFKKFVVSVFVIALVVAISTGLIVNIVI*

15 Sequence 2347

Contig_0744_pos_3530_4705,
 is similar to (with p-value 2.0e-71)
 >sp:sp|P39643|AAT2_BACSU PROBABLE ASPARTATE AMINOTRANSFERASE
 (EC 2.6.1.1) (TRANSAMINASE A) (ASPART). >pir:pir|S39740|S397
 20 40 hypothetical protein - Bacillus subtilis >gp:gp|X73124|BS
 GENR_86 B.subtilis genomic region (325 to 333). NID: g413923
 . >gp:gp|Z99123|BSUB0020_65 Bacillus subtilis complete genom
 e (section 20 of 21): from 3798401 to 4010550. NID: g2636240

25 atgaggaggcatagcatggaaatgtcagaaaggctagcttcaattcctgatagctacttt
 ggcaaaacaatgggcccgtatagttgaacatggtcctttaccacttataaatatggcagtt
 ggaattccagatggagaaacgcaaaagggtattatcaatcatttttcagaggcgctatgt
 attccagaaaaatcaaaagtatggtccatttcacggcaaaagatgcctttaacaagctatt
 gtttaacttctaccaaagacattacgactgttgtaattagacaaagaagatgaagttgtatt
 30 ttatatgggactaaaaatggtcctgttgctacacttctgtgttgaatcctggtgaa
 attgtacttttacctgatccgggatatacagattatttagcgggggtcatgttagctgat
 gctaagccactccctttaaaattgtcgccaccaaattatttgcgaattggaatactata
 agtgctaaagttccttgagaagactaagctaatttatttaacatatcccaataatcctacc
 ggcttcgacagcgacacaagatgatatttgatgaagcgattcatcgttttaaggtactcaa
 35 acaagatagttcatgaactttgcatatagtgcttttgattgacgcaaaaatccaagc
 atattagcttctaaaaatgcaaaagatggtgctatcgagatattctctttatctaaaggt
 tataatatgtcaggccttctgtgtgggtttgctgttggttaataaaaaaatgattcaagcg
 ttaaagaagtatcaaaactatacaaatgcaggatggtttggagcattcaagatgtgct
 acgtatgcactcaatcattatgatgagtttttagaaaagcaaaatgaaatatttagacgt
 40 agacgtgataattttgaatcacaactaaaacatgcacatttaccggtttgttcaacttaag
 ggaggtatttacatttgggttacatacaccgccgggttatgatagtgaaagcattcgaacag
 ttgttatataaagaaaagtcatttttagttgcacctggtaaacatttgggtgaaaatggt
 aatcaatatgtgagggtttcattggcgctcgatgataaacaattagaagaagcggcgaat
 cgcttaacacaattacggtatttgtatgaaagataa

45

Sequence 2348

MRRHSMEMSERLASIPDSYFGKTMGRIVEHGPLPLINMAVGIPDGETPKGIINHFSEALC
 IPENQKYGPFHGKDAFKQAIWNFYQRHYDVELDKEDVCILYGTKNGLVALPTCVVNPGE
 IVLLPDPGYTDYLAGVMLADAKPLPLKLSPPNYLPWNNTISAKVLEKTKLIYLYTPNNPT
 50 GSTATQDDFDEAIHRFKGTQTKIVHDFAYSAGFGDAKNPSILASKNAKDVAIEIFSLSKG
 YNMSGFRVGFVGNKKMIQALKKYQTHTNAGMFGALQDAATYALNHYDEFLEKQNEIFRR
 RRDNFESQLKHAHLFPVHSGGIIYIWLHTPPGYDSEAFEQLLLKEKSILVAPGKPFGENG
 NQYVRVSLALDDKQLEEAANRLTQLRYLYER*

55 Sequence 2349

Contig_0744_pos_4718_5716,
 is similar to (with p-value 7.0e-84)
 >gp:gp|U31175|SAU31175_1 Staphylococcus aureus D-specific D-
 2-hydroxyacid dehydrogenase (ddh) gene, complete cds. NID: g

1644432.

atgactaaaattaaattaatgggtgtcagagaagaagatgaacattatattgaaatgtgg
 tcacaacaacatgaagtggagtgatgtcgaaagaacagttaactgaagacaatgtc
 caatctattgaaggatttgatggactatcattgtctcaaacattaccattatcagaaaca
 5 atttataataaattaaatcaacttgggaattcggcagatcgctcaacgaagtgtggattt
 gatgggtataatttagagtttagcatctaaatatggtcttattatatctaatgtgccttcc
 tattcacctcgaagcattgctgagtttaccgtgactcaagccatcaatattgtacgtcac
 tttactcatattcaaagaaaaatgagattgcacgatttttaggtgggaagcatcaatttta
 tctcaatcaatcaaagatttaaaggtagcggttattggcacgggacatattggtggcatt
 10 gttgcacaaatattctcagaaggatatctatgtgacgtttagcgatgatccttttcca
 agtgaacatgtgaaaccttacgttacctataaacaagataaatgaggcaattaaagag
 gcagatattgtcacaatacatatgccgtcaacacataataacaattacctgtttaatgaa
 aacatgtttcaaattgtttaaaaagggtgctgtgtttgtaaattgtgctagaggatcctta
 gtagataccaaggtttgttatctgcaatagagcaaggtaaataaagggtgcagcactt
 15 gatactttatgaattgaaattggagtatatatcgacagatagaagtgaaaggtttgaat
 gaccacttttagagggaatttaattactagagaagatattattgttacaccgcataatagca
 ttttatactgaaggaggaatcaaacatcttatttttgatgcttttagatgcaacaatggaa
 gtattaaatactggcacgacggagttaagagtaaattaa

20 Sequence 2350

MTKIKLMGVREEDEHYIEMWSQQHEVEVDMSKEQLTEDNVQSIIEGFDGLSLSQLPLSET
 IYNKLNQLGIRQIAQRSAGFDGYNLELASKYGLIISNVPSYSPRSIAEFTVTQAINIVRH
 FNHIQRKMRLHDFRWEASILSOSIKDLKVAVIGTGHIGGIVAQIFSEGYLCDVVAYDPFP
 SEHVKPYVYTKQSINEAIKEADIVTIHMPSTQYNNYLFNENMFQMFKKGAVFVNCARGSL
 25 VDTKALLSAIEQQQIKGAALDITYEYIGVYTTDRSEEGNDPLLEELITREDIIVTPHIA
 FYTEEAIKHLIFDALDATMEVLNTGTTEL RVN*

Sequence 2351

Contig_0744_pos_8456_6108,
 30 is similar to (with p-value 0.0e+00)
 >sp:sp|P32113|ATKA_ENTFA POTASSIUM/COPPER-TRANSPORTING ATPAS
 E A (EC 3.6.1.36). >pir:pir|A45995|A45995 Cu2+-transporting
 ATPase (EC 3.6.1.-) - Enterococcus hirae >gp:gp|L13292|ENECO
 PPUMP_1 Enterococcus hirae ATPase (copA) gene, complete cds;
 35 ATPase (copB) gene, complete cds. NID: g290641.
 atgacatgtgctgctgctcaaatcgattgaaaagaaattgaatcgatgaatcatggt
 caagctaaagtgaatctgactactgaaaaagcaactatcgactatgagctgacgattat
 catctcgaagattttgtagagcaaattcaaagctcggctatgatgttgagttgagcaa
 gtagaattaaatataaatgggtatgacatgtgctgcatgttctaatacgatagaaaaggt
 40 ctaaatcaaacgcaaggtgtacaacaggcaacagtaaaacttaactaccgaacaagcactc
 atcaaatattaccctagtgtctacgaacacggaagcattaattaagcgtattcaaaatatt
 ggatacgatgctgaaactaaaacttcataaaagcgcaatcaaatcgtaaaaaacaagag
 ttaaaacataaaacgcaataaattaatcatttcagctattttatcggtgccactattatta
 gtaatgggtggtgcatatctcacctatttccattccatccattttggtcaatccttgggtg
 45 caattaattctttcaacacctgtccaatttattattgggtggcaattttacgttggcgcg
 tataaaaatttgcgaaatgggttcagctaacaatggatgtattgggtgctgttggtaccagt
 gccgcatatttttatagcatttatgaaatgatgatgtggctcacacatcaaacacatcac
 ccgcatattatatttgaacaagtgtctattttaattacgttaattcttcttggttaaatat
 ttagaagcagctgcaaatcaaacactacacagactaccaatgcattaaagcgaattgttaaat
 50 gcgaaagaagcagcagtaattaaagaaaaataagaaattatgcttccacttgataaagtt
 aaagtcggagatactttactaataaaaacccggcgaaaagatacctgtagatggcaaagtc
 actaaaggtgatacttctattgacgaatccatgctaactggtagtctatcctgttgaa
 aaaagtagtggcgattcagtgattggttctaccatgaataaaaatggttcaatcatgatt
 gaagcaactcaagtaggtggtgatactgcattatcacatataattaaagtggttgaggat
 55 gctcaaagttctaaagcaccgattcaacgcttagctgatattatttctggatattttgtt
 ccgattgttagttagcattgctggttattacttttatcatatggattatattcggttcccc
 ggcaatttgaacctgcactgtttcagcaatatctgttttagttattgcttgcctgt
 gcacttgggttagcaacgcctacatctattatggtaggtacaggacgtgctgcagaaaat
 ggcatattattcaaaggaggccaattttagaagcgtgcacattatggtgatacaatcgtg

5 ctagataaaacaggcacaattactaatgggtcaacctgtagtaactgattatgttggtgac
 aatgatacattacaacttttagcaagtgtgaaaatgcttcagaacatcctcttgctgat
 gctattgttacttatgctaaagataaaggtcttaatttacttgataatgacacttttaa
 tcaattccgggacatggtattaaagctacgattcatcaacaacaaatccttggggcaat
 10 cgaaaattaatgaacgattacaatatatctattagtaataaataaatgaccaatataat
 cactatgaacatttaggtcaaacggcaatgatgattgccgtggataatcaaatatgga
 atcattgctgttgctgatacagtaaaaaatgatgctaacaagcgataaaaagaactaaga
 aatatgaatatcgactggttatgctgactggtgataacaatcgacagctcaaaccatc
 gccaaacaagttggcattgaacatgtaattgcagaagtgtgccgaagaaaaggcacat
 15 caaatctctttattacaagacaaaggtaaacaggttgccatggtcggtgatggaattaat
 gatgcgcctgcacttgtaaaagccgatattggaatggctataggcactggagctgaggta
 gcgattgaagctgcagatattacgattccttggtggtgacttgctattagttccaaaagct
 atcaaaagcaagtaaaagctacgattaaaaatattcgacaaaattttttgggcatgtgga
 tataacgtagctggcatcccaatagctgcttggtggtttattagcaccttgattgcccgt
 20 gctgctatggcattaagtctctgtagcgtagttatgaatgcattaagactgaaaaaatg
 aaactatag

Sequence 2352

20 MTCAAQSNRIEKKLRNMNHVQAKVNLTTTEKATIDYESDDYHLEDFVEQIQSLGYDVAVEQ
 VELNINGMTCAACSRIEKLNLQOTQGVQATVNLTTTEQALIKYPSATNTEALIKRIQNI
 GYDAETKTSSKAQSNRRKQELKHKRNLIIISAILSLPLLLVMVHISPISIPILVNPVW
 QLILSTPVQFIIGWQFYVYAYKNLRNGSANMDVLVAVGTSAAFYFYSIYEMMMWLTHQTHH
 PHLYFETSAILITLILLGKYLEARAKSQTTNALSELLNLQAKEARVIKENKEIMLPDKV
 25 KVGDTLLIKPGEKIPVDGKVTGKDTSIDESMLTGESIPVEKSSGDSVIGSTMNKNKSGIMI
 EATQVGGDTALSHIIKVVEDAQSSKAPIQRLADIISGYFVPIVVSIIVITFIWIIIFVHP
 GQFEPALVSAISVLVIACPCALGLATPTSIMVGTGRAENGILFKGGQFVERAHYVDITV
 LDKTGTITNGQPVVTDYVGDNDTLQLLASAENASEHPLADAIPTYAKDKGLNLLDNDTFK
 SIPGHGKATIHQQQILVGNRKLMDYNISISNKLNDQLNHYEHLGQTAMMIAVONQING
 IIAVADTVKNDKQAIKELRNMNIDVVMVLTGDNRTAQTIKQVGIHVIAEVLPEEKAH
 30 QISLLQDKGKQVAMVGDGINDAPALVKADIGMAIGTGAEVAIEAADITILGGDLLLVPKA
 IKASKATIKNIRQNLFWAFGYNVAGIPIAACGLLAPWIAGAAMALSSVSVMNALRLKLM
 KL*

Sequence 2353

35 Contig_0745_pos_341_742,
 is similar to (with p-value 5.0e-21)
 >sp:sp|P35154|YPUG_BACSU HYPOTHETICAL 29.6 KD PROTEIN IN RIB
 T- Γ ACB INTERGENIC REGION (ORFX7). >pir:pir|S45549|S45549 hyp
 othetical protein X7 - Bacillus subtilis >gp:gp|L09228|FACDI
 40 A_16 Bacillus subtilis spoVA to serA region. NID: g410114. >
 gp:gp|299116|BSUB0013_34 Bacillus subtilis complete genome (section 13 of 21): from 2395261 to 2613730. NID: g2634723.
 atgtattcatgctataataatttagtgaaatgcatgaacgatgcaatttgatatagag
 gtagatgtaattgatgaagtaaaactcgatgcatttaattggtccattagacttattattg
 45 catctaattcaaaaatgatgaattgatatttatgatatccctatgaaagccttaactgaa
 cagtacatgcaatatgttcagtcgcatgaatcagctagaaattaatgttgctagtgaatat
 ttagttatggcatcagaattactaatgattaaaagtaaatcaacgatgaatataaccaa
 aatccaacactaaaaacatgctgtgcggtctctgcactatgaactgttggttcaccatt
 aactgttgtagtaaaattagcatccatcgctccagtaattga
 50

Sequence 2354

MYSCYNLKVYALTMQFDIEVDVMEYVKLDAFNGPLDLLHLIQKYEIDIYDIPMKALTE
 QYMQYVHAMNQLEINVASEYLVMASELLMIKSKFNDEYNQKSNTKNMLCGFCTMNCCFTI
 55 NCCSKISIHRSN*

Sequence 2355

Contig_0745_pos_1717_2316,
 putative peptide of unknown function
 gtgacttrtatgtctacgaataatgagattgaatttaaacaaatactagatcaagatact

tactcaaaaatctatgaacactatttcaaaaatcaatcaccttttaagcaaactaatttc
 tatatcgacacagagaatttttaaatataaacagcatcatgctgctttgcgataagggtg
 aaagattatatgtttgaaatgactttaaaagtcctgctgaagttggattgacagaatat
 aatcactcagtaaatatagaacctgaacttgatgctcacttcaactttctcaattaccc
 5 aacgataattagaatatatttagaacaggactttaatatatttagaaaatgagcttaaagta
 ctaggaaacttaactacctatcgtttagaaaccgattatcaaaatgaattactagtatta
 gataagagtgaatatctcggcaaaactgattatgaattagagtttgaagttcattcttat
 gatgaaggataattcaaaatttaaaactttacttcaacatttttaattctcaacatcaaaaa
 cccttgaataaagtgcaacgtttttttcaagaaaaacaaaatgcaagtataaagagtaa
 10

Sequence 2356

VTFMSTNNEIEFKQILDQDTYSKIYEHYFKNQSPFKQTNFYIDTENFKLKQHHALRIRV
 KDYMFEMTLKVP AEVGLTEYNHNSVNIEPELDMSLQLSOLPNDIRNILEQDFNILENELKV
 15 LGNLTTYRLETDYQNELLVLDKSEYLGKTDYELEFEVHSYDEGYSKFKTLLQHFNLQHOK
 PLNKVQRFFQEKQNASDKE*

Sequence 2357

Contig_0745_pos_2380_2763,
 20 putative peptide of unknown function
 atggaacacgggtgatattatgtctaaaacaccatagagttgattggtcaaaaagccttg
 tatcaaatgattgatcattttctatcaacttgctcgagaaagattctcgatcaatcattta
 ttcccagcgattttcaaggaaaccagtcgaaagcagaagcaatttttgacacagtttctt
 ggaggtcctgacttatataaccaagaacatggtcatcccatgttaaaacgaagacatatg
 25 gaatttacaattagcgagtgatgaacgtgatgcatggcttgagaacatgcatactgctatt
 caacacgccaacttcctgcggtgtaggcgattacttgtttgagcgattaagacttact
 gcaaatcacatggtaaattcctaa

Sequence 2358

MEI:GDIMSKTPYELIGQKALYQMIDHFYQLVEKDSRINHLFPGDFKETS RKQKQFLTQFL
 30 GGPDLTYQEHGHPMLKRRHMEFTISEYERDAWLENMHTAIQHAKLPAGVG DYLFERLRLT
 ANH MVNS*

Sequence 2359

Contig_0745_pos_2807_3583,
 35 putative peptide of unknown function
 atggagaataagagtcgtgaagataactatcacctgttagcaaaaatagaaatctat
 tctttttttgatccttttagcaaaagattgttttaattatctgcaatcttatcaaaatta
 agaattgaatataataaatatataaaggtaagacatattttaaaccttcttttaaaggta
 40 ttaactaagtgtcaagctcaaagtaacttcagattttgacaatattgcacttgctataaa
 gccgctgaacttcaaggtcgtatcagagcagaaagatttatacatttaattgcaaaatgaa
 atcattccaaaacgtgatattattaccgaagatatgatttctgattgtattaataatgcc
 ggcattgactatcaagtttttaaagaagacttgcaaaaggacaagttgactgacagcttg
 aaagttgatcttcacattgcaagagaaatggaaatagaacaagctccctcacttgttttt
 45 ttcagcgaatgttcatgaagaagggtttaaaagtcgaaggattatccttatcatatt
 tatacttacattattaatgagtttaattgggacaacctatagagaaaaatcttcctccaaaa
 ttagaatactacattcaaaagaacaactagtaacaatggaagaacttttaacgatctat
 gaatggcctgaaaaattgctaataaagaattaaagaaactcacacttcaacaaaaagtt
 gaaaagttgcaatatccagagggagaatttttgaaatctaaaatgcctcagtgtta
 50

Sequence 2360

MENKSREDTNLSPVSKIEIYSFFDPFSKDCFKLSAILS KLRIEYNKYIKVRHILNPSLV
 LTKCQAQSTSDFDNIALAYKAAELQGRIRAERFIHLMQNEIIPKRDIITEDMISDCINNA
 GIDYQVFKEDLQDKLTDLSKVDLHAREMEIEQAPSLVFFSENVHEGLKVEGLYPYHI
 55 YTYIINELMGQPIEKNLPPKLEYIYQKKQLVTMEELLTIYEWPEKLLNKELKKLTLQKKV
 EKLQYPEGEFWKSKMPQC*

Sequence 2361

Contig_0745_pos_5858_4050,

is similar to (with p-value 0.0e+00)
 >gp:gp|088209|D88209_1 *Bacillus licheniformis* DNA for Pz-pep-
 tidase, complete cds. NID: g1651215.
 atgagtcacaattaacaagagaagaacaggaacgtaaatatcctgaatatacatgggat
 5 ttaacaactatTTTTTAAAAGTGATGAAGCATTGGAAGAAGCTTTAAAAGTATTGAAGCT
 AAAATAGGTGAAGAAGAAAAATTTAAAGGTCATCTTGGTGAAAGTGCTGAAACATTATAT
 GAAGCGCTAAGTCTTGAAGACGAGTTAGGTACAAAATTAGAAAAGGTATATGTATACGCA
 CATTTAAAACAAGATCAAGATACTGCAATGATAAATATACCGGTTTAGAAGCGCGTGCA
 CATCAACTTGTATTAAATATAGCTCTGCATGGAGTTTTTTAGTACCTGAAATTTTACAA
 10 CTAGATGAAGCTACTATTCAATCTTTATCGATTCTAATGATGATTTAAAACGATATGAA
 TTCGATTTGAAATTGATTAATGAGAAGCTCCACATATATTAGATGCGAATACAGAAAAG
 TTATTAACAGAAGCACAAGACGCCTTTCAACGCCTTCTAATGTATATGGAATGTTTCAGC
 AATGCAGATTTAGAATTTGAAGATGCTATAGATAAAGATGGTCAAGCTTATCCTTTAACA
 CAAGGTACATTTATCAAGTATTTAGAGTCTGATGATCGTGAGTTAAGAGCTTCTGCTTTT
 15 AGAAATGTTTATAAAGCATACGGTGCGCATAATAACACGCTAGGTGCTACTTTAGCTGGT
 GAGGTTAAGAAAAATGTATTAAATGCTAGAAGTCATCATTATCGTTCAGCACGTGAAAGA
 GCTTTAAGTAATAATCATATTCAGAAGCTGTTTACGATAACTTAATCAAAACGGTCCAT
 AAATACTTACCTTTATTACACAGATACACGAAGCTTAGACAAGAGTTACTAGGTTTAGAC
 GATTTAAAAATGTATGATCTTTATACACCTCTTGTAAAGATGTCAAATTTGAAATGCCA
 20 TATGAAGAGGCAAAATCCTGGATGTTAAAAGCACTTGAGCCAATGGGAGAAGAATACTTA
 AACGTGGTTAAGGAAGGTCTAGATAACCGTTGGGTCGATGTATATGAAAATAAAGGTA
 CGTTCAAGYCGGATATTCATCCGGTGGACATTTAACTAATCCTTTTCACTTTACTTAAGT
 TCAGACACTGTTTCTGATTTATATACTTTAGTACATGAATTTGGTCACTCTGCACATAGT
 TACTTTAGTAGACAGAATCAACCATCAAATTTAAGCGATTATACAATCTTTGTGCTGAG
 25 GTAGCATCAACTTGTAATGAGGCTTACTTAGTGACTACATGGACAAACATTTAGATGAT
 GAACGACGCTCTATTGTTACTTAACCAAGAATTAGAAGCATTAGAGCAACACTATTCCTG
 CAAACAATGTTTGCTGAATTTGAACATAAAATACATCAAATAGAAGAAGCTGGGGAGCCG
 TTAACGCCAAATCGTATGAATGAAGAATATGCTAAACTGAACAACTATATTTGGTGAA
 GCAGTAGAAACTGCAGATGATATTAGTAAAGAATGGTCACGTATTCCTCATTCTATATG
 30 AATTATTATGTATATCAATACGCAACTGGTTATAGTGACGCTCAAAGTTAAGTCATCAA
 ATTTTAACTGAGGGTCAACCTGCTGTTGAACGATATATCAATGAATCTTAAAAAAGGGT
 AGCTCAAACCTATCCGATTGAAATTTAAAAAATGCAGGTGTTGACATGACAACACCTCAA
 CCAATAGAGGAAGCTTGTGAAGTATTCGAACAAAAATTAGATGCTTTTGAAGGTTAATG
 35 AAGCTTAG

Sequence 2362
 MSQQLTREEQERKYPEYTWDLTTIFKSDEAFEEAFKSIEAKIGEEEFKKGHLGESAEITLY
 EALSLEDELGTKLEKVYVYVYHLKQDQDTANDKYTGLEARAHQLVIKYSSAWSFLVPEILQ
 40 LDEATIQSFIDSNDLKRFEFDLKLINERPHILDANTEKLLTEAQDALSTPSNVYGMFS
 NADLEFEDAIDKDGQAYPLTQGTFFIKYLESDDRELRAAFNRNVYKAYGAHNNTLGATLAG
 EVIKNVFNARTHHYRSARERALSNNHIPEAVYDNLIKTVHKYLP LLHRYTKLRQELGLD
 DLKMYDLTPLVKDVKFEMPYEEAKSWMLKALEPMGEEYLNVVKEGLDNRWVDVYENKKG
 RSGGYSSGGHLTNPFILLNWSDTVSDLYTLVHEFGHSAHSYFSRQNPNSLSDYTI FVAE
 VASTCNEALLSDYMDKHLDDERRLLLLNQELERFRATLFRQTMFAEFHKKIHQIEEAGEP
 45 LTPNRMNEEYAKLNKLYFGEAVETDDDISKEWSRI PHFYMNYYVYQYATGYSAAQSLSHQ
 ILTEGQPAVERYINEFLKKGSSNYPIELKNAGVDMTTPQPIEEACEVFEQKLDAFEKLM
 KA*

Sequence 2363
 50 Contig_0745_pos_1454_846,
 is similar to (with p-value 2.0e-77)
 >sp:sp|O06446|SECA_STAAU PREPROTEIN TRANSLOCASE SECA SUBUNIT
 . >gp:gp|U97062|SAU97062_1 *Staphylococcus aureus* NCTC 8325 S
 ecA (secA) gene, complete cds. NID: g2078389.
 55 atgggtgggtattgctatacataaaggtgatattgcagaaatgagaacaggtgaagggaaa
 acattgactgcaaccatgccgacgtatttgaatgctttagctggtagaggtgtacatggt
 attacagtcgaatgaatatctatcaagttcacaaaagtgaagaaatggctgaactatataac
 tatcttggcttaactgtaggtttgaacttaaatagtaagtcgaactgaagaaaaacgtgag
 gcttacgcacaagatatcacttatagtagcaataatgaacttgggtttgattatcttaga

gataatatggtgaactatgctgaagagagagtaatgcgtcctctacattttgcaattatt
 gatgaggtcgattccatattgatcgacgaagcaagaacaccttaattatttctggtgaa
 gcggaaaaatctacttctttatatggaaaattgattattattagttataaattattcaaa
 aataaagagcaactattttatgttgcataatgaaataaagggaagaaagtctactataaa
 5 cgagatatataatcctaaagattatattaaagaacataaacctgattgtaaggatatacgt
 agaaaatga

Sequence 2364

10 MGGIAIHKGDIAEMRTGEGKTLTATMPTYLNALAGRGVHVITVNEYLSSSQSEEMAELYN
 YLGLTVGLNLNSKSTEEKREAYAQDITYSTNNELGFDYLRDNMVNYAEERVMRPLHFII
 DEVDSILIDEARTPLIISGEAEKSTSLYGLIIISYKLFKNKEQLFYVAYEIKGKKVYYK
 RDINPKDYIKEHKPDCKDIRRK*

Sequence 2365

15 Contig_0745_pos_0_300,
 putative peptide of unknown function
 atggattattattcaatttcataagcatcaacattacttccctatgtcatgatatttta
 gaagatgcttggaatcacaaaatactttttcaaacatgatgcagtagtaagtctcatt
 ttgtttgcaacagctaattatcactatcgctcggtggaatttaaaaggcgcatataaatct
 20 tttgaaaaagctttattaactattaaccacgctaaagatgacactcaattacaccttaat
 atcaatgaatttaaacacttattgtgaaaatgattgaagcagtaaaattacaaaaaACT

Sequence 2366

25 MDYYYQFHKHQHYFLCHDILEDWKSQNTFSKHDAVVSILFATANYHYRRGNLKGALKS
 FEKALLTINHAKDDTQLHLNINEFKQLIVKMIEAVKLQKT

Sequence 2367

30 Contig_0746_pos_541_1500,
 is similar to (with p-value 0.0e+00)
 >sp:sp|P05425|ATKB_ENTFA POTASSIUM/COPPER-TRANSPORTING ATPAS
 E B (EC 3.6.1.36). >pir:pir|B45995|B45995 Cu2+-transporting
 ATPase (EC 3.6.1.-) - Enterococcus hirae >gp:gp|L13292|FNECO
 PPUMP_2 Enterococcus hirae ATPase (copA) gene, complete cds;
 35 ATPase (copB) gene, complete cds. NID: g290641.
 atgcatcatgataaccatgcctcacatcatatagtgccatgcacatcatcatggaat
 tttaaagttaagttttttgtttcattaatttttgcataacctatcattctcttatcgcca
 atgatgggtgttaacttaccttttcaattcacatttccaggttctgaatgggtagtgtta
 atattaagtacaattttattcttttatggtggtaaacgggttctgtctggtggtaagat
 40 gaaattgctacaaaaaaaccagcatgatgaccttagttgctctaggtatttcagtagct
 tatatttatagcttgtatgctttttatatgaataacttttagtagtgcaactggcataca
 atggacttttttgggaattagcaaccttaatttttaattatgctattaggacattggata
 gaaatgaatgctgctggaaatgctggagatgctttaagaaaatggcagaactattacct
 aatagtgtctattaaagttatggataatggccaacgcgaagaagttaaaatatcagacatc
 45 atgactgatgatatcgctgaagtaaaagccggagaaagcattccaacagatggtattatc
 gttcaaggacaaacatctatagatgaatccctagtcactggagaatctaaaaaagtacaa
 aaaaatcaaaatgacaacgtcatcgggggttctattaatgggtctggaacaatacaagtc
 aaggttacagctgttggagaagatggatatctttctcaagttatgggacttgttaatcaa
 gcacaaaaatgataaatctagtgtgaattgttatctgataaagtagcgggttatttattc
 50 tactttgctgtaagtgttgccgtgatttcttttattgtctggatgctcattcaaaatgat
 gttgattttgcattagaacgtcttgaactgtgttagtcattgcttgcctatgcttttag

Sequence 2368

55 MHHDNHASHHHSGFLAHGHNFKVKFFVSLIFAIPIILLSMMGVNLPFQFTFPGSEWVVL
 ILSTILFFYGGKPFLLSGGKDEIATKKPGMMTLVALGISVAYIYSLYAFYMNNFSSATGHT
 MDFFWELATLILIMLLGHWIEMNAVGNAGDALKKMAELLNSAIKVMNDNGQREEVKISDI
 MTDDIVEVKAGESIPTDGIIVQGQTSIDESLVTGESKKVQKNQNDNVIGGSINGSGTIQV
 KVTAVGEDGYLSQVMGLVNQAQNDKSSAELLSDKVAGYLFYFAVSVGVISFIVWMLIQND

VDFALERLVTVLVVIACPML*

Sequence 2369

Contig_0746_pos_1614_2543,

- 5 is similar to (with p-value 2.0e-78)
 >sp:sp|P05425|ATKB_ENTFA POTASSIUM/COPPER-TRANSPORTING ATPAS
 E B (EC 3.6.1.36). >pir:pir|B45995|B45995 Cu2+-transporting
 ATPase (EC 3.6.1.-) - Enterococcus hirae >gp:gp|L13292|ENECO
 PPUMP_2 Enterococcus hirae ATPase (copA) gene, complete cds;
 10 ATPase (copB) gene, complete cds. NID: g290641.
 atggataaaaactggtactttaactgagggtaacttttctgtgaatcattatgagagcttt
 aaaaatgatttgagtaatgatacaatattaagccttttcgcctcattagaaagtcaatct
 aatcaccatttagctataagtattgttgattttgcgaaaagtaaaaaatgtttcatttact
 aatccacaagacggttaataatattccagggtgtcggattagaaggtctaattgataataaa
 15 acatataaaaataacaaatgtctcttattcttgataaacataaaacttaattatgacgatgac
 ttgtttactaaattagctcaacaaggtaattcaatcagttattttaattgaggatcaacaa
 gtcattggcatgattgtctcaaggagatcaaattaaagaaagctcaaaacaaatgatagct
 gatttactatcaagaaatattacaccagtcagtgcttacagggtgacaataatgaagtggca
 cagctgtcgcgaaaagaattaggtattagtgatgttcacgcacaactcatgccagaagat
 20 aaggaaagcattataaaagattatcaaagtacggttaataaagtcagtgatggtcggagac
 ggtatcaacgatgcgcgagctcttataagagccgatattggtatagcaattggtgcaggc
 acagatgttgacgtggattcaggtgatatcatacttgtttaaagtaaatccatcagatatc
 attcatttcttgactctttcaataataactatgagaaaaatggtgcaaaacttatggtgg
 ggtgcagggtataatattgttgctgtaccttttagcagctggcgcattagcttttatcggg
 25 ttaalattatcaccagctgtaggagcaatattaatgtctttaagtacagttatagtagcg
 attaagtcttttacattaaaattaaaataa

Sequence 2370

- MDKTGTLTEGNFSVNHYESFKNDLSNDTILSLFASLESQSNHPLAISIVDFAKSKNVST
 30 NPQDVNNIPGVGLEGLIDNKTYKITNVSYLDKHKLNYYDDDLFTKLAQQGNSISYLIEDQQ
 VIGMIAQGDQIKESSKQMIADLLSRNITPVMLTGDNNEVAHAVAKELGISDVHAQLMPED
 KESIIKDYQSDGNKVMVVDGINDAPSLIRADIGIATGAGTDVAVDSGDIILVKSNPDI
 IHFTLSNNTMRKMVQNLWWGAGYNIVAVPLAAGALAFI GLILSPA VGAILMSLSTVIVA
 INAF TLKLK*

35

Sequence 2371

Contig_0746_pos_4011_4493,

putative peptide of unknown function

- 40 atgatgaaaaaagataaagacactaatgacaaaaaagtgagagccatatgaagcataat
 gatgaaagtaaaagttcctgaagatatgacatcgactaatgaggggtgaatttaaagtagga
 gataaagtaacgattacagcaggacatatgccagggtatgaaaggtgcagaagctactgta
 aaaggtgcgtataaaacatatgcctatggttgtaagttataaaccacaaatggaaatgaa
 aaagtaaacatcataaatggatcgtaaacgaagaaatcaaagatgcacctaagatgga
 tti:agtaaaggcgatactgttaaattagaagcaagtcatatgtctggtatgaaaggtgct
 45 acagccaatatagataacgtgaaaaagacgactgtttacgtagttgattacaaatccaaa
 gataatggtaaaatcattaaaaatcataaatggatgacaggaaatgagctgaaagcacga
 taa

Sequence 2372

- 50 MMKKDKDNTDQKSESHMKHNDESKVPEDMTSTNEGEFKVGDKVTITAGHMPGMKGAEATV
 KGAYKTYAYVVSYKPTNGNEKVNHNKWIIVNEEIKDAPKDGFSKGDVVKLEASHMSGMKGA
 TANIDNVKKT TVYVVVDYKSKDNGKIIKNHKWMTGNELKAR*

Sequence 2373

55 Contig_0746_pos_5109_4705,

is similar to (with p-value 8.0e-65)

>sp:sp|P30330|ARSC_STAAU ARSENATE REDUCTASE (ARSENICAL PUMP
 MODIFIER). >pir:pir|D41903|D41903 arsenate reductase (EC 1.-
 .-.)- Staphylococcus aureus plasmid pI258 >gp:gp|M86824|PI

2ARSRBC_4 Plasmid pI258 arsenic resistance operon (arsRBC) genes, complete cds. NID: g150725.

gtgtcaataatgacaaagaaaataatttactttatatgtacaggtaattcatgccgtagc
 caaatggctgaaggctggggcaaaaacatcttaggtgatgaatggcaagtatatctggc
 5 ggtattgaggcacacgggtgtcaatccaaaggcaattgaagcaatgaaagaagttggaatt
 gatattctcaaacatacttctaatttaattgacaacactataactaaaccaatccgattta
 gtagtaactctatgtagtgtatgctgacaataattgtcctatattaccgccgaatgtaaaa
 aaagaacattggggatttgatgatccagcaggcaaaccttggtcagaattccaacgtgtt
 agagatgaaatcaaaaccgccattgaatcatttaaaactagataa

10

Sequence 2374

VSIMTKKIIYFICTGNSCRQMAEGWGKNILGDEWQVYSGGIEAHGVNPKAIEAMKEVGI
 DISNHTSNLIDNTILNQSDLVVTLCSADANNCPILPPNVKKEHWGFDDPAGKPWFSEFQRV
 RDEIKTAIESFKTR*

15

Sequence 2375

Contig_0746_pos_4270_3908,
 putative peptide of unknown function

atgattgtttactttttcatttccatttgggtttataacttacaacataggcatatgt
 20 tttatacgcaccttttacagtagcttctgcacctttcatactggcatatgtcctgctgt
 aatcgttactttatctctacttttaaatccacctcattagtcgatgtcatatcttcagg
 aactttactttcatcattatgcttcataatggctctcacttttttggtcattagtgtcttt
 atcttttttcatcatattaaaatctccctattgattagtttggttacttttatttgacca
 tcattccattatcttcatgctcaagtatatgacagtgaacatgtatgttccagtatttt
 25 taa

Sequence 2376

MIVYFFISICGFITYNIGICFIRTFYSSFCTFHTWHMSSCNRYFISYFKFTLISRCHIFR
 NFTFIIMLHMALTFVLVISVFIFHHIKISLLISLLLLFDPSFHYLHAQVYDSETCMFQYF
 *

30

Sequence 2377

Contig_0747_pos_4050_4487,
 is similar to (with p-value 8.0e-27)

35 >sp:sp|P42435|NASD_BACSU NITRITE REDUCTASE (NAD(P)H) (EC 1.6
 .6.4). >gp:gp|D30689|BACNARB_4 Bacillus subtilis DNA around
 narB region (nasB operon and nasA gene). NID: g710016. >gp:g
 p|Z99105|BSUB0002_159 Bacillus subtilis complete genome (sec
 tion 2 of 21): from 194651 to 415810. NID: g2632457. >gp:gp|
 40 D50453|D50453_33 Bacillus subtilis DNA for 25-36 degree regi
 on containing the amyE-srfA region, complete cds. NID: g1805
 369.

atgttatggagggtttgttatgggaagtttttttaatcggtgactcgaaaagagaatcct
 actgtagagtctggtgttaaagattttggcgatcatatctgttgaaaatggctaccaaata
 45 tttatcgagggtaatggtggtactgatgttactgtaggtaaattgttaacgacagttgaa
 accgaagatgaagtgattcaattatgtggtgccctcatgcagtattacagagaaacaggt
 gtttacgctgaaagaacagcaccatggttagaacgtatgggctttgaaaatgtcaagaat
 gtcttattaaatcaagaaaagcaaaaagaactgtatttaagaattatggaagccaaaaaa
 gctgttgagaatgaaccatgggaaactattgttgaaaataaagaagcacaaaaaatcttt
 50 gaagttgagaaggtgtaa

Sequence 2378

MLWRFVMSFFNRMTRKENPTVESGVKDFGVISVENGYQIFIGNGGTDVTVGKLLTVE
 TEDEVQLCGALMQYYRETGVYAERTAPWLERMGFENVKNVLLNQEKQKELYLRIMEAKK
 55 AVENEPWETIVENKEAQKIFEVEKV*

Sequence 2379

Contig_0747_pos_4490_4804,
 is similar to (with p-value 3.0e-22)

- >sp:sp|P42436|NASE_BACSU ASSIMILATORY NITRITE REDUCTASE (NAD (P)H) SMALL SUBUNIT (EC 1.6.6.4). >gp:gp|D30689|BACNARB_5 Bacillus subtilis DNA around narB region (nasB operon and nasA gene). NID: g710016. >gp:gp|299105|BSUB0002_158 Bacillus subtilis complete genome (section 2 of 21): from 194651 to 415810. NID: g2632457. >gp:gp|D50453|D50453_32 Bacillus subtilis DNA for 25-36 degree region containing the amyE-srfA region, complete cds. NID: g1805369.
- 5 atgaaagctaaagaaaagattaaagttacaacaatgaatgaaatgattcctcaaataggc
 10 aaaaaagtagttgtaaacgaaaaagaaataggtatcttctcacagataatggtgattta
 tatgccattggaaatatatgtccacataaagaaggaccgttgtctgaagggaactgtaagt
 ggtgattatgtttactgtccgttacacgatcaaaaaatagctttaaaaactggagaagta
 caacaacctgatacaggatgtgtagagacatacgaagtagaagttattgatggagatatt
 tacttatgtctataa
- 15 Sequence 2380
 MKAKEKIKVTTMNEIPQIGKKVVNEKEIGIFLTDNGDLYAIGNICPHKEGPLEGTVS
 GDYVYCPHLDQKIALKTGEVQPPDTGCVETYEVEVIDGDIYLCL*
- 20 Sequence 2381
 Contig_0747_pos_3285_2647,
 is similar to (with p-value 8.0e-40)
 >sp:sp|P39592|YWBI_BACSU HYPOTHETICAL TRANSCRIPTIONAL REGULA
 TOR IN EPR-GALK INTERGENIC REGION. >pir:pir|S39679|S39679 hy
 25 pothetical protein - Bacillus subtilis >gp:gp|X73124|BSGENR_25
 B.subtilis genomic region (325 to 333). NID: g413923. >gp:
 :gp|299123|BSUB0020_126 Bacillus subtilis complete genome (s
 ection 20 of 21): from 3798401 to 4010550. NID: g2636240.
- 30 atgaatgacattgtgaacgttcaaaaagggtcatattaaaataggcttatcaccatgatg
 aatgttcaaatgtttacaatgcattgaatcagtttcacagactctatcctaattgtgaca
 tatgaagtgtattgaggtgtgtggttaaaattgttgagaacttaacatcctaattgatgtg
 gatattggtattactacattacctgtagatcacactgaatttcattcaacttctttatat
 aatgaagaattattattagtagtaagtaattgaccatcatttagcacatttaataaagta
 gacatggcagatttgaaagatgaagagtttgtttttttcatgatgattattatttaaaa
 35 gatcaaatatagagaactgtaaaaggctaggctattaccctaaaactgttgctaattat
 tctcaaattagttttatcgctaataatgattcaacaaggaataggaattagtatcggtcca
 gaaagttagtttaatttaattggtggaataacgtaacgtccattcaattagagaatgttgaa
 ttatcatggcatcttggcgtgatattggagaaaagatgcttatctcaatcatgtaactcgc
 aaatggattgaattttttctgagatgaaaccaacatag
- 40 Sequence 2382
 MNDIVNVQKGHKIGLSPMMNVQMFTNALNQFHRLYPNVTYEVIEGGGKIVENLTSNDDV
 DIGITTLVPDHTFEFSTSLYNEELLVVSNDHHLAHLNKKVDMADLKDEEFVLFHDDYYLK
 DQIIENCKRLGYYPKTVANISQISFIANMIQQGIGISIVPESLVNLMGNVTSIQLENVE
 45 LSWHLGVIWRKDAYLNHVTRKWIEFISEMKPT*
- Sequence 2383
 Contig_0747_pos_1831_863,
 is similar to (with p-value 8.0e-89)
 50 >pir:pir|A25805|A25805 L-lactate dehydrogenase (EC 1.1.1.27)
 - Bacillus subtilis
- 55 atgaaggagttcgttaaaatgaaaaaatttgggaaaaaagttgttttagtaggagacggt
 tccgtaggttcaagttatgcatttgctatggtgactcaaggaattgcagatgaatttgta
 attattgatattgcaaaagataaagtggagcagacgttaaagatttaaacatggtgca
 ctttacagttcttcaccagtgactgtaaaagctggagaatatgaagattgtaaaagatgca
 gatttagttgttattacagcaggtgcacctcaaaaaccgggtgaaactcgtttacaactt
 gttgagaaaaataactaaaatcatgaaaagtatcgtaactagtgtcatggatagtggttt
 gatggtttcttctcctaattgtctgcaaacccagttgatattctaacacgttatgttaaagaa
 gttacaggtttaccagctgaacgtgttattggttctggtacagtgcttgatagtgcaaga

ttcagatattttaataagtaaagaattaggtgttacatcaagtagtggtcacgctagcatt
 ataggtgaacatgggtgactctgaacttgcaagtttggtctcaagcaaactgtggaggtatt
 tcagtgtatgatacatgaaagaagaactggtagcgatgctaagcgaaatgaaatttat
 attaatacaagagatgctgcttacgatattcattcaagctaaaggatctacgtattatggt
 5 atagctctagcactattacgtattttctaaagctttactaaataatgaaaatagtattttg
 acagtttctagtcacctaataatgggtcaatatggatttaacgatgtttatcttggttacca
 acacttatcaatcaaaaatgggtgcagttaaaatttatgaaacaccattaaatgataacgaa
 ctacaattactagaaaaatcagtgaaaactttagaagacacttatgattctataaaacat
 ttagtttaa

10

Sequence 2384

MKEFVKMKKFKGKKVVLVGDSVGSYAFAMVTQGIADFEVVIDIAKDKVEADVVDLNHGA
 LYSSSPVTVKAGEYEDCKDADLVVITAGAPQKPGETRLQLVEKNTKIMKSI VTSVMSDGF
 DGEFLIAANPVDILTRYVKEVTGLPAERVIGSGTVLDSARFRYLISKELGVTSSSVHASI
 15 IGEHGDESELAVWSQANVGGISVYDTLKEETGSDAKANEIYINTRDAAYDIQAKGSTYYG
 IALALLRISKALLNNENSILTVSSQLNGQYGFNDVYLG LPTLINQNGAVKIYETPLNDNE
 LQLEKSVKLTLED TYDSIKHLV*

Sequence 2385

20 Contig_0747_pos_800_450,
 is similar to (with p-value 1.0e-34)
 >gp:gp|L16975|LACALS_2 Lactococcus lactis alpha-acetolactate
 synthase (als) gene, complete cds. NID: g473900. >gp:gp|A23
 961|A23961_1 L. lactis alpha-acetolactate synthase gene. NID
 25 : g809617.

atggcggaaaaacaatatctgcagcacaaaatggtaattgatactttaaaaataatgga
 gttgagtagtatttgggtattccaggtgcgaaaatcgactacttatttaaatgcactagag
 gatgacgatattgaattagtcggttacgcgtcatgaacaaaacgcagcgatgattgcacaa
 ggtattggctggttaacaggaaaaccaggtgtggctattactacaagtggcccaggggta
 30 agtaacttaactactggtttattaactgcaacttctgaaggtgaccctgtattagctatc
 ggtgggtcaagttaaaaagaaatgacttattacgtttaacatcaatacgcctaa

Sequence 2386

MAEKQYSAAQMVIDTLKNNGVEYVFGIPGAKIDYLFNALEDDDIELVVTRHEQNAAMIAQ
 35 GIGRLTGKPGVAITTS GPGVSNLT TGLLTATSEGD PVLAIGGQVKRNDLLRLTSIR*

Sequence 2387

Contig_0748_pos_1246_3342,
 is similar to (with p-value 0.0e+00)
 40 >sp:sp|P52026|DPO1_BACST DNA POLYMERASE I (EC 2.7.7.7) (POL
 I). >gp:gp|L42111|BACPOL_1 Bacillus stearothermophilus DNA p
 olymerase I (pol) gene, complete cds. NID: g806280.

atgaaaggtctaattggggatacctctgacaatatctctggcggttgctggtgtcggcgaa
 aagacggctattaaattacttaataatcttgagtcagtagaaggggtctatgaacatatt
 45 gagggaggtcactgcaaaaaaattaaaaagaaaaactcatcaatagtaaagatgatgcctta
 atgagtaaagatttagcaacaatcaatgttcacagtcaggattgaagtatcattagaagat
 acaaaattaactctacaagacgacactacagaaaaaattgaactatttaaaaagctagaa
 tttaaacaactatttagcagatatagacacatcctctacgaatgaagaagtcatagataaa
 acttttgaaattgagcaagactttcaaaatgtagatttgatgatttaaacgaagcggta
 50 atacattttgaactcgaaggcactaattatcttaagacactattctcaagtttggtttt
 tatacaataatcacaatgtagtgataaatgctgaggatgtaaggattataaacattta
 gttcaatggccttgagataaaaaatacaactaaaattgtctatgatgcaaaaaaaacttat
 gtatctgctcatcgattagggttaatatagaaaatattgaatttgatgttatgtagca
 agctatattattgaccatcacgttctattgatgacgttaaatctgttgtaagtttatat
 55 ggacaaaattatgtaaaagataatattacaatatgtgggaaaggtaagaaacatcatata
 cctgaagaaccaattctaaacgaacacattgcctctgtgacagaagctatagcagctgta
 actccaaccattgaatcacagttagaagattataatcaaatgaactgttgaaagattta
 gaatttaacattgaacagaatttttaagtgaattggaagaaattgggtatataaccggtatc
 aatgatttgaaagaaatggaatttcgaaattcaaaaaaattggatgtattaatatccaat

attcatgagtcggctggtgaagcgtttaatatcaattctcctaagcaattaggtgttgtt
 ttatttgaacattacaattgcctgtcattaagaagacaaaacgggctattcaacagct
 gtagacgtattagaaaaactacaaggtgagcatcctattatagatgatattttagaatat
 agacaacttgctaagttgcaatctacgtatgtagagggattacaaaaagtaataagcaaa
 5 gatcatagaattcacacacggttttaatacaaacgcttgctcaaaactggtagattatcaagt
 atagatcctaatttacaaaaatatacctatacgattagaagaaggaagaaagattagaag
 gccttttaaccaacttctaaagatagtgtgattttatctgctgattattcacaattgag
 ttacgtgtacttgctcatattacgcaagatgaaagtttaaacatgcatttataaacgga
 catgatattcacactgcaacagcaatgaaagtatttaatgttgaatctgaccaggttgat
 10 agtttaatgagacgtcaagcaaaagctgttaactttgggtattgtatatgggtatcaocgat
 tatggattgagtcagagcttggtattactagaaaaacaagcaaaagcatttattgatgat
 tatttagctagttttccaggtgtaaaacaatatatgtcagacattgttaagatgcaaaa
 gcacaaggttatgtggaacactacttcatcgctcgatcacattcctgatataacaagt
 agaaacgttaatttaagaagttttgcagaaagaacagcaatgaatacacccatacaaggt
 15 agtgcagctgacataataaaaattagcaatggttaaattcagtgaagaaagattaaagaaact
 aaatatcatgctaagttattattacaagttcatgatgaactcatatttgaataacaaaa
 tcagaagtagaagtttttagtaaatgttagaagaaattatggaacaagcattagtgtc
 gatgtacctttaaaagtagattcgaattatgggtgcaacatggtacgatgctaaataa

20 Sequence 2388
 MKGLMGDTSNIPGVAGVGEKTAIKLLNQFESVEGVYEHIEEVTAKKLKEKLINSKDDAL
 MSKDLATINVHSPIEVSLEDTKLTLODDTTEKIELFKKLEFKQLLADIDTSSTNEEVIDK
 TFEIEQDFQNVLDNLNEAVIHFELEGTNYLKDTILKFGFYTNHQHVINAEDVKDYKHL
 VQWLEDKNTTKIVYDAKKTYVSAHRLGINIENIEFDVMLASYIIDPSRSIDDVKSVSLSY
 25 GQNYVKDNITIFGKGGKHHIPEEPILNEHIAVTEAIAAVTPMKSQLEDYNQIELLKDL
 ELPLARILSEMEEIGIYTDINDLKEMEFIEIQKKLDVLISNIHESAGEAFNINSFKQLGVV
 LFETLQLPVIKKTGTGYSTAVDVLEKLQGEHPIIDDILEYRQLAKLQSTYVEGLQKVISK
 DHRIHTRFNQTLAQTGRLLSSIDPNLQNIPIRLEEGRKIRKAFKPTSKDSVILSADYSQIE
 LRVLAHITQDES LKHAFINGHDIHTATAMKVFNVESDQVDSLRRQAKAVNFGIVYGISD
 30 YGLSQSLGITRKQAKAFIDDLASFPGVKQYMSDIVKDAKQGYVETLLHRRRYIEDITS
 RNVNLRSAERTAMNTPIQGSAADIKLLAMVKFSEKIKETKYHAKLLLQVHDELIFEIPK
 SEVEDFSKFVEEIMEQALVLDVPLKVDSNYGATWYDAK*

Sequence 2389
 35 Contig_0748_pos_4255_4857,
 is similar to (with p-value 2.0e-31)
 >sp:sp|Q55515|Y553_SYNY3_HYPOTHETICAL 22.5 KD PROTEIN SLR055
 3. >gp:gp|D64006|SYCSLLLH_95 Synechocystis sp. PCC6803 compl
 ete genome, 25/27, 3138604-3270709. NID: g1001291.
 40 gtgattgggataactggtggtattgccactggaaaatcaacagtttcagaattattaaca
 gcatatgggttttaaaatcgtagatgctgatattgcttcacgcgaagcagttaaaaaaggc
 tctaagggtcttgaacaagttaaagagatttttggggaagaagcaattgacgaaaatggt
 gagatgaatcgctcaatatgtaggagagatagtttttaatcatcctgacttacgcgaggct
 cttaatgaaatagttcatcctattgtaagagagataatggaacaagagaaaaacaattat
 45 ctagaacatggatatcatgtaattatggatatcccattgttgtagcaaaaatgaactacaa
 gatactgtagatgaagtttgggtggtttatacatctgaaagtattcaaatcgatcggttta
 atggagaggaataatttatcattagaagatgctaaagcacgtgtttatagtcaaatatct
 atagataaaaaaagtaggatggcagatcatgtgatagataatctaggtgataaattagaa
 cttaaacagaattttacaaaaattacttgaagaagaagggtatattcaatcggagag;gaa
 50 tag

Sequence 2390
 VIGITGGIATGKSTVSELLTAYGFKIVDADIASREAVKKGSKGLEQVKEIFGEEAIDENG
 EMNRQYVGEIVFNHPDLREALNEIVHPIVREIMEQEKNNYLEHGYHVIMDIPLLYENELQ
 55 DTVDEVVVVYTSESIQIDRLMERNNLSLEDAKARVYSQISIDKKS RMADHVIDNLGDKLE
 LKQNLQKLLEEGYIQSESE*

Sequence 2391
 Contig_0751_pos_2197_2541,

putative peptide of unknown function

atggagatgatagaagaacgtaatttatcagggttattcaaactaactttcatcat
 cccatcattcaaattcttaagagaacacattaaatcaacttaaaatactctctcattat
 ttaccagagcgacacctgcaatgggtggcaattcaatcttggtcacaatggtttactgat
 5 catgggattactgaaatccaccttgatgtaactgcacaagcgcttagatcttattacaaa
 ggtattttttataaaatgtcatcttataaaatactgctcatagcgttttgacaggtggatat
 tatcacggttcactagaagggttttggttttaggattaacacttta

Sequence 2392

10 MEMIEERNLSGLIQTLTFNHPIIQILKENTLNQLKILSHYLPERHPAMVAIQSWSQWFTD
 HGITEIHLDVDTAQAPRSYYKGIFIKCHLKNTAHSVLTGGYYHGSLEGFGLGLTL*

Sequence 2393

Contig_0751_pos_2566_3180,

15 is similar to (with p-value 5.0e-27)

>sp:sp|Q02129|HIS1_LACLA ATP PHOSPHORIBOSYLTRANSFERASE (EC 2
 .4.2.17). >pir:pir|D45734|D45734 HisG - Lactococcus lactis s
 ubsp. lactis >gp:gp|U92974|LLU92974_4 Lactococcus lactis unk
 nown gene, partial cds, and HisC (hisC), unknown, HisG (hisG)
 20), unknown, HisB (hisB), unknown, HisH (hisH), HisA (hisA),
 HisF (hisF), HisIE (hisIE), unknown, unknown, LeuA (leuA), L
 euB (leuB), LeuC (leuC), LeuD (leuD), unknown, IlvD (ilvD),
 IlvB (ilvB), IlvN, IlvC (ilvC), IlvA (ilvA), AldB (aldB) and
 aldR (aldR) genes, complete cds. NID: g2565137.

25 atgttacgagttgcattagcaaagggtcggtttattaaagagttttatcgaatatttaca
 caagttaatcagatagatattgcaactgtacttttaatatagacagcgacagttattgctt
 acagtcgacaacattgaaatgatlttagttaaaggaagcgatgtgcctacttatgtagaa
 caaggtattgctgatgtaggaatagtgggaagtgatattctgaatggtcaaaaatataat
 attaataaattactcgatttgccatttggtaaatgtcattttgcttggtggcgcaagcca
 30 gaaacatctcgctataaaaaagtagcaacaagctatgtacatacagctactcaattcttt
 aataaagaaggtatggatgtagaagtgattcaccttaacggttcagttgaattgtcatgt
 gtagtggatagggtgatgctattgtagatattgtacaaactggttctacgcttacagct
 aacgggctcggttgaaaaagcatatcagtgaattaacgctaagttaattacaaataaa
 gaatcatattttaagcaatcatctgaaatagagagactaatcaagcagttaggagtgct
 35 attaactatgcttag

Sequence 2394

MLRVALAKGRLLKSFIEYLQOVNQIDIATVLLNRQRQLLLTVDNIEMLVKGSDVPTYVE
 QGIADVIGVSDILNGQKYNINKLLDLPGKCHFALAAKPETSRYKKVATSYVHTATQFF
 40 NKEGMDVEVIHLNGSVELSCVDMVDAIVDIVQTGSTLTANGLVEKKHISEINAKLITNK
 ESYFKQSSEIERLIKQLGVSINYA*

Sequence 2395

Contig_0751_pos_3320_4003,

45 is similar to (with p-value 2.0e-41)

>sp:sp|Q02136|HISX_LACLA HISTIDINOL DEHYDROGENASE (EC 1.1.1.
 23) (HDH). >pir:pir|E45734|E45734 HisD - Lactococcus lactis
 subsp. lactis

gtggaaactgagaagcttgaattagagcaaagccaactaaaaaatgcatacgacatgcta
 50 gataatgaaacacgagatgcattagagcaaagctatcagagaattaaagtgtaccaagaa
 aatattaaggttaaaacaggaatcatctcaaaaactgaatggttatgaacgataccatcct
 atcgaacgtgtaggattttatgtgccgggaggttaaggctagctatccgtctacagtatta
 atgactgcaacacttgctcaagtagcaggtgttaatgagattactgttggtacccacct
 caaaatagcgggtatatgtcaagaggtgttagccgcttggtacattacaggcggttcacat
 55 gtttatcaagtcggtggagcacaaagtattgcggcgctaacttatggcacggaaactata
 aaaaaagtcgacaaaatcgtaggtccagggaatcaatatgttgcttatgcaaaaaagttt
 gtattcggtcaagtaggcataagacaaatcgacagaccgacagaaatagccttgattata
 gacgaaagtgtgacttagacgcaatcgcttatgacgtatttgcaacagcagaacatgat
 gaaatggcttgacttatgtgattagtgaaaatgaaaaggtacttaataatgaacact

ataatacaagagaaacttcagtag

Sequence 2396

5 VETEKLELEQSQLKNAYDMLDNETRDALEQSYQRIKQYQENIKVKQESSQQTCEYERYHP
IERVGIYVPGGKASYPSTVLMATLAQVAGVNEITVVTPPQNSGICQEVLAACYITGVHH
VYQVGGAQSIAALTYGTETIKKVDKIVGPGNQIVAYAKKFVFGQVGDQIAEPTETIALII
DESADLDAIAYDVFAQAEHDEMACTYVISENEKVLNQLNTIIQEKLQ*

Sequence 2397

10 Contig_0751_pos_832_338,
is similar to (with p-value 3.0e-78)
>sp:sp|P51065|PPCK_STAAU PHOSPHOENOLPYRUVATE CARBOXYKINASE (
ATP) (EC 4.1.1.49). >gp:gp|U51133|SAU51133_1 Staphylococcus
aureus phosphoenolpyruvate carboxykinase (pcka) gene, comple
15 te cds. NID: g1255261. >gp:gp|L42943|STAPEPCK_1 Staphylococ
us aureus (clone KIN50) phosphoenolpyruvate carboxykinase (p
cka) gene, complete cds. NID: g860731.
atgtatcatttcttaagtggattcacgtctaaactagctggaacagaacgtggtgttact
gaacctcaaccttcggttttcaacttgctttggtgcaccattcttacctttgagtccaaca
20 aagtacgctgatctacttgaaatttaacgatattcatgatgtatgtatcttagta
aatactggatggacaggtggttaaatatggtgtagggcgaagaattagtctacactatact
cgtgaaatggtagatcaagcaatatcaggtaaatataaaaaactaaatatattaaagat
gatacatttggtctaaatattccagttcaaattgacagtgtagctacaactattctgaat
cctatcaatgcttggaacaataaagataactacaaagcacaagcttacgatttgattcaa
25 cgctttaataataatttttaaaaaatttcggcaaggaagtcgaacatatattgccacaaaggt
gcatttaatacaataa

Sequence 2398

30 MYHFLSGFTSKLAGTERGVTEPQPSFSTCFGAPFLPLSPTKYADLLGNLIDIHDVDVYL
NTGWTGGKYGVGRRISLHYTREMVDQAIISGKLKNTKYIKDDTFGLNIPVQIDSVPTILN
PINAWNKNKDYKAQAYDLIQRFNNNFKFGKEVEHIANKGAFNQ*

Sequence 2399

35 Contig_0752_pos_2931_2140,
putative peptide of unknown function
atgagcaataacatgtggagggggagaaagttgaacttaaaattagaccatatcattcac
tatttacatcaattagagtcatttaagtttcccgagaaatattagaattgcaaaatggt
ggaagacatcatcatttgggcacctttaatcaaatagcaccgattaaaaatagttatc
gaattgctagatggtgaaaatgagtcaaaacttagcaatatagctaaaactgaagaaggt
40 cgtgtatcatttgcataaaaaatagtcaggatcattttaacaaggggttaaaggtatt
tgttttagaacaaaggatataaatcaggttaaaagtacttttagaaaaatagaggcgttgat
gtgataggctcctattgatatggaagagaaaaacaaaaaaggtcatcaaattcgttggaga
ttgctatatattgctaaccctgactatacagtcacaaaccctttctttatagaatgggat
aacaacaaaaagcaaaacctatcacaataacataatttcaacttgcatcggtttaaaatt
45 aaagaggtgatttattactagcactcaacgtgaaacaacagtaagtccttggaaagaatgg
tataacctgaaaatagtaaatgaaacggctacatctactgatctcaaattagaaactgat
gaagttatctataaaatagaagacggcaagattcaggttttcataattaataatgacc
gatatcaatgccacagcaccatattcaatatttatcgtggtgctaaatatcggtttgag
50 ccaccaactag

Sequence 2400

MSIHNIWVRCKRLNLKLDHIIHYLHQLSFKFPGEILELQNGRRHHHLGTFNQIAPIKNSYI
ELLDVENESKLSNIAKTEEGRVSFATKIVQDHFQGVKGICFRTKDINQVKSTLENRGVD
VIGPIDMERENKKGHQIRWRLLYIANPDYTVKPPFFIEWDNNKKQNLSQLHNFNLSSFKI
55 KEVITSTQRETTVSLWKEWYNLKIVNETATSTDLKLETDEVIYKIEDGKDSGFHTLIMT
DINATAPYSIFIRGAKYRFEPNP*

Sequence 2401

Contig_0752_pos_1740_490,

is similar to (with p-value 3.0e-24)

>gp:gp|AF099966|AF099966_1 Staphylococcus sciuri factor essential for methicillin resistance FEMA (femA) gene, complete cds. NID: g3820631.

5 atggaaaagatgaacatcactaatcaacaacatgacgcatttgtgaaatctcatcccaat
ggtgattttattacaattatctaagtgggcagatacgaaaaattaacaggatggtattca
agaagaattgctgtcgggtgaaaatggtcaaattaaagggtgttgccagctactattcaaa
aaaatacctaacttccatacactttatgctatgtatctaggggattttagctgattat
aataataaagaagtgtagaagctctacttagctatgctaaaagaagtagcaaaagatgaa
10 aagtcgtatgctatcaaaatagatcccgatgtcgaagtagataaagggtgcagaagcactt
aaaaatctacgtgagcttgggttttaacataaagggttttaagaaggactgtctaaagac
tatattcaaccaagaatgactatgattacgcctattgacaaaacagatgatgaattagtt
caaagtttcgaacgtcgaaatcgttcaaaagtaagacttgactgaagcgtggaactaaa
gtagaacgatcaaatcgcgagggggttaaaatcttctgaatttaataagataaactggg
15 gagagagatgggtttttaaactcgagataattagttattttgaaaatatatatgatgcactt
catgaagacggtgatgcagaactcttcttgttaaatagagcctaagccagtattagat
acggttaatacaagatcttgaagcacaattagctgagaaagagaaattacaatcaaaaaag
caagataaaaagacacttaataaaacttaataatgatattgataataaaattaagaaaacaaat
gaattaaaatcggatttaacagaacttgaaaaaagcgagccagaaggattttacttgtca
20 ggagcgtcttaattgtttgcaggaaaacaaatcttactatctctatggcgcttctcgaat
gactatcgtgatttcttaccaaaccatcacatgcaatttgaaatgatgaaatatgcacgt
gagcatgggtgcaacaacctatgactttgggtggtacagataatgatcctgataaagattca
gaacattatgggttgtgggcttttaaacgagtttgggtacatatttaagtgaaaaaatt
ggagaatttgattatgtattaaatcaaccgctatatcatttagttgagaaagtgaacct
25 cgtttaacgaaagctaaaattaaaatatcacgtaaacttaagggtaaataa

Sequence 2402

MEKMNITNQQHDFAVFKSHPNGDLLQLSKWADTKKLTGWYSRRIAVGENGQIKGVGQLLFK
KIPKLPYTLCYVSRGFVADYNNKEVLEALLSYAKEVAKDEKSYAIKIDPDVEVDKGAEAL
30 KNLRELGFKHKGFKEGLSKDYIQPRMTMITPIDKTDELVSFERRNRSKVRLALKRGTK
VERSNREGLKIFANLMKITGERDGFTRDISYFENIYDALHEDGDAELFLVKLEPKPVLD
TVNQDLEAQLAEKEKLQSKKQDKKTLNKLNDIDNKIKKTNELKSDLTELEKSEPEGIYLS
GALLMFAGNKSYYLYGASSNDYRDFLPNHMQFEMMKYAREHGATTYDFGGTDNDPKDS
EHYGLWAFKRVWGTYLSEKIGEFDYVLNQPLYHLVEKVKPRLTKAKIKISRKLKKG*

Sequence 2403

Contig_0752_pos_0_359,
putative peptide of unknown function

40 atggtagttttaattatattaggtggcgtttattcaagcgccaaattaaaacttgaatta
ttaccagatggtgaaaatccagttatttcagttcaaaactacaatgtctggagcaacaccc
cagtcacaacaagatgaaataagtagcaagattgataatcaagtagcgtcgttggcctac
gtaaatagtgtagcagactgaatctatacctaatacttctatagtaactgtagaatacgtat
aatggtagagatatggataaagctgaagaacaattaaaaaagaaatcgacaaaattaaag
45 ttttaagatggcgttgggtgaaccgaattaacaaggaactctatggatgctttTGTTT

Sequence 2404

MVLLIILGGVYSSAKLKLLELLPDVENPVISVQTTMSGATPQSTQDEISSKIDNQVRLAY
VNSVQTESIPNASIVTVEYDNGTDMKAEELKKEIDKIKFKDGVGEPELTRNSMDAFVX

Sequence 2405

Contig_0753_pos_2155_2514,
putative peptide of unknown function

55 atgataaatattatattgaagaaaatagacttgaggttaattagaatgttcgttgttaca
aatagaatcactgtaaaaaaggatattgcaaaacaaatggcgctaattttactaaagga
ggacctattgaaatctttaaagggtttgaaggattgaagtttgcaaatgataaagat
gattatagcgaagatatgtatgtaaatagttggtgggaaactgaagaagattttaaaaat
tgggtgaatagtgtattttaacaagcacataaaaaactggaaaatccgaagattca
ccagtcattaaaagcgaaattgttaaatcaaatgttttatcttcttgaacagaagataa

Sequence 2406

MINIILKKIDLEVIRMFVVTNRITVKKGYAKQMAPNFTKGGPIESLKGFEIEVWQIDKD
 5 DYSEDMYVNSWWETEEDFKNWNVNSDVFKQAHKNTGKSEDSPIVKSSEIVKSNVLSLNR*

Sequence 2407

Contig_0753_pos_4019_2700,
 10 is similar to (with p-value 3.0e-88)
 >sp:sp|P30267|YKAA_BACFI HYPOTHETICAL 50.9 KD PROTEIN IN KAT
 A 3'REGION (ORF A). >pir:pir|S27491|S27491 hypothetical prot
 ein A - Bacillus firmus >gp:gp|L02548|BACKATA2_1 B.firmus OR
 F A and ORF B, complete cds. NID: g143118.
 15 atgtcat*gttagggattatattatTTTTtagttcctataaccagtcggttcaagatg/jaaag
 cagcaaacgacacttcctatagctTTTTtagctggTTTTattaaaagattggcttggtggt
 atcatgccaaattTTtaattgtaaccatcataactgtatcaggtattTTtaacaatattatgc
 tctacaattTTataaaaaataaattaaatcctcaaggTTtaatgagcagtgctTTcaacgTT
 aaaataggatggcttgTTTTgagagtattagctgtcttctTTcttggttaacattTTta
 20 aatattggacctgaaatgattaaatctgaagatacaggtggattagtagttttcaagTTta
 ttacctactcttgtagcagtagtttttatttgctgcaatctTTttacctttattaatggag
 tatggtctattagaattacttggacctatcttagacctatcatgcgacctttgtttact
 ttacctggtagatcgacagttgataatctagcttcatTTtataggtgatggtacagttggt
 gTTTTaattactagtagacaatatggtgaaggatattactctagaagagaagcaacagta
 25 atatccacaacctTTtagtggtgtatctattacggttcgctattgtcattgccgaaacaatt
 agaatgcaagatcaattTTtctattTTtttaaacagttgtcatttcatgcttaattgca
 gcaatgattatgccaaagaatttgccaccttaaaaatattcctgacgaatatgctaaagaa
 gtaagtgaagaggctcgtaatgaacagctaccagaaggcaaaacagcattaaaatatggt
 tttgatttagcaactgaagttggaattaaatcgccagggtTTtaagaattTTtaatttca
 30 ggTTTTaaaacagttgtagatatgtggtttgtaattttaccagttggtatgagtatagga
 acaatagctaccattattgctaaactacacgcctgTTTTtgaaattataggaaaaccattt
 gttccagtactagaattgttacaaattccagaagcacatgaagcatcacaacaattTTta
 attgggtttgccgatattgttcttaccttcaattcttattgaaggggttcaaatgatgta
 acacgttttgtaattggagcattgagtagtctcacaactgtgtatttatctgaagtggc
 35 ggctgattcttgggttctaaaattccagttagtataagtaaattatttatgatttttta
 attcgtagctatcattacgcttccaataattgctttattagcgcatttatttatcgjataa

Sequence 2408

40 MSLLGIILFLVPIPVVQDGKQQTTLPIAFLAGLLKDWLGGIMPILIVTIITVSGILTILC
 STIYKKNLNPQGLMSSAFNVKIGWLVLRVLAVFFSWLTFNLNIGPEMIKSEDTGGLVFSS
 LPTLVAVFLFAAIFLPLMEYGLLELLGPIFRPIMRPLFTLPGRSTVDNLASFIDGTVG
 VLITSRQYGEgySRREATVISTFVSVITFAIVIAETIRMQDQFFYFYLTVVISCLIA
 AMIMPRIWPLKNIPDEYAKEVSEEARNEQLPEGKTALKYGFDLATEVGIKSPGFKEFLIS
 45 GFKTVVDMWFVILPVVMSIGTIATIIANYTPVFEIIGKPFVPVLELLQIPEAHEASQTIL
 IGFADMFLPSILIEGVQNDVTRFVIGALSISQLVYLSEVGGVILGSKIIPVSISKLFMIFL
 IRTIITLPIIALLAHLFIG*

Sequence 2409

50 Contig_0753_pos_1617_1108,
 putative peptide of unknown function
 atgacaggtaaaacacacgcatcatgtggctTTTTtagtcggtgcaataaccacacaatat
 tttcatcacagatatatttacttctatatcagtgattgtactttcagtcatttcaagtata
 ttgccagatatatgtcatacacaaagtaaaataggaagacgatttaggcttactag:ttt
 55 tttgtcagaattttatttggtcatagaacatttacgcattcacttttatttatta:agga
 attagttttttactgtacttcatacaaaactccgatgtattatatgggtgcaattgttatt
 ggtatgttttcgatgtttatcttgatatattaaccccaagagggtgtaaaactatttatat
 cttttaccatttaataatcgtagtaccattcatttttaaaactgggggactagtagatgta
 tctctagctactgcattaagtgttggtgcgatataactttatttcaaccatattttaa

actatgatgcactattggttaatcaaataa

Sequence 2410

5 MTGKTHASCGFLVGAITTQYFHTDIFTSISVIVLSVISSILPDICHTQSKIGRRFRLTSF
FVRILFGHRTFTHSLLFIIGISFLLYFIQTPMYMVAIVIGMFSHVILDILTGRGVKLLY
PLPFNIVSPIHFKTGGLVDVSLATALSVGAIYTLFQPYLNTMMHYWLIK*

Sequence 2411

10 Contig_0754_pos_2035_2466,
putative peptide of unknown function
atgatacaaggttttaggctatttattgtccaatataacagattataaagaattaacgaat
ttagctcaaaatggagatcgtgatgccattgatttaaaagtaaaacatatttataaaagat
actgaaccaccaatttcctggagatttaacagcagcaaattttggaaatgtattacatcac
ttagataatcagtttacatcagctaacaacttgccctgcaattggcgctggttggtgaa
15 gttataacaactatggctattacattagcacgtgaatataagactaagcacgttgatatat
atcggttcatcatttaataacaatcaattactacgtgaagttgttgaaaattacactgtt
ctaagaggatttaaaccgtactatattgagaatggtgctttttcaggcgctttaggagca
ctttacctctaa

20 Sequence 2412

MIQGLGYLLSNITDYKELTNLAQNGDRDAIDLKVKHIYKDTEPPIPGDLTAANFGNVLHH
LDNQFTSANKLASAIGVVGEVITTMATLAREYKTKHVYIGSSFNNQLLREVVENYTV
LRGFKPYYIENGAFSGALGALYL*

25 Sequence 2413

Contig_0754_pos_4500_5048,
putative peptide of unknown function
atggatgattataaagaatatagaaaaagacttatcggttaaattaaaaaacctataggt
agagatttatataatagattatataaaaaatattcaagatactttagaacctgaggtttat
30 gaaattgctcctaataactaaattaggacattttcctggttatcagaatgtaacggttatcc
caccacaaaatgcaacaaattatatcaagaaatgaacctagttgaaacaagctttaatg
aatgtaaaaggtgtttatgttataaccgacttaagtaatggcaaatatacataggtatca
gcatcaggtaatactgatggaatatggcaacgatggtcggactatgccaacatagaaaaat
ctaacaggtggttaataaattattaaatgaaattaaattagataaaagggaaagattacatc
35 ataaataattttcaatattcaatttttagagatttttgatacaaagactaaggtagacact
ataatcaatagagaaaattattggaagaatgtattttgcactagaaaatattggtatgaac
tttaactaa

Sequence 2414

40 MDDYKEYRKRLIVKLLKKPIGRDLYNRLYKNIQDTLEPEVYEIAPNTKLGHFPGYQNVTL
HPQMQQIISRNEPSWKQALMNVKGVYVITDLSNGKLYIGSASGNTDGIWQRWSDYANIEN
LTGKNKLLNEIKLDKGKDYIINNFOYSILEIFDTKTKVDTIINRENYWKNVFCTRKYGMN
FN*

45 Sequence 2415

Contig_0754_pos_5848_7851,
is similar to (with p-value 0.0e+00)
>sp:sp|P05425|ATKB_ENTFA POTASSIUM/COPPER-TRANSPORTING ATPAS
E B (EC 3.6.1.36). >pir:pir|B45995|B45995 Cu2+-transporting
50 ATPase (EC 3.6.1.-) - Enterococcus hirae >gp:gp|L13292|ENECO
PPUMP_2 Enterococcus hirae ATPase (copA) gene, complete cds;
ATPase (copB) gene, complete cds. NID: g290641.
atgcatcatgataacctgcctcacatcatcatagtgccatgcacatcatcatggaat
tttaaagttaagttttttgcttcatatttttgcaatacctatcattctcttatcgcca
55 atgatgggtggttaacttacctttttcaattcacattttccaggttctgaatgggtagtgta
atattaagtacaattttattcttttatggtggttaaaccgttctgtctggtggttaaagat
gaaattgctacaaaaaaaccaggcatgatgaccttagttgctctaggtatttcagtagct
tatattttatagcttgatgcttttttatatgaataacttttagtagtgcaactgggtcataca
atggactttttttgggaattagcaaccttaattttaattatgctattaggacattggata

gaaatgaatgctgtcggaatgctggagatgctttaagaaaatggcagaactattacct
 aatagtgctattaaagttatggataatggccaacgcgaagaagttaaaatcagacatc
 atgactgatgatcgtcgaaagtaaagccggagaaagcattccaacagatggtattatc
 gttcaaggacaaacatctatagatgaatccctagtcactggagaatctaaaaaagtacaa
 5 aaaaatcaaaatgacaacgtcatcggggttctattaatgggtctggaacaatacaagtc
 aaggttacagctgttggagaagatggatatctttctcaagttatgggacttgtaatacaa
 gcacaaaaatgataaatctagtgctgaattgttatctgataaaagtagcgggttatttattc
 tactttgctgtaagtggtgctgatttcttttattgtctggatgctcattcaaaatgat
 gttgattttgcattagaacgtcttgtaactgtgtagtcattgcttgtccacatgcttta
 10 ggcttggcaataccttttagtcactgcacgttctacttcaattgggtgcacataatgggtta
 attattaaaaatagagagctctgtagaaatagctcaacatatcgattatgtaatgatggat
 aaaactggtaactttaactgagggtaacttttctgtgaatcattatgagagctttaaaaat
 gatttgagttaatgatacaaatattaagccttttcgctcattagaaagtcaatctaatacac
 ccattagctataagttattgtgattttgcaaaaagtaaaaatgtttcatttactaatcca
 15 caagacgttaataatattccaggtgtcggattagaaggtctaattgataataaaacatat
 aaaataacaaatgtctcttatcttgataaacataaacttaattatgacgatgacttggtt
 actaaattagctcaacaaggttaattcaatcagttatttaattgaggatcaacaagtcatt
 ggcattgaltgctcaaggagatcaaattaaagaaagctcaaaacaaatgatagctgttta
 ctatcaagaataattacaccagtcattacaggtgacaataatgaagtggcacacgct
 20 gtcgcaaaaagattaggtatttagtgatgttccacgcacaaactcatgccagaagataaggaa
 agcattataaaagattatcaaagtgcggttaataaagtcattgatggcggagacggtatc
 aacgatgcgcgagtgcttataagagccgatattggtatagcaattgggtgcaggcacagat
 gttgcagtggttcaggtgatatacattgttaaaagtaattccatcagatatcattcat
 ttcttgactctttcaaaataatactatgagaaaaatgggtgcaaaacttatgggtgggtgca
 25 ggttataatattgttgctgtaccttttagcagctggcgcattagcttttatcggttaata
 ttatccacagctgtaggagcaatattaatgtctttaagtacagttatagtagcgattaat
 gcttttacattaaaattaaaataa

Sequence 2416

30 MHHDNHHASHHHSGHAHHHGNFKVKFFVSLIFAIPILLSPMMGVNLPFQFTFPGSEWVVL
 ILSTILFFYGGKPFSLSGGKDEIATKKPGMMLVALGISVAYIYSLYAFYMNFFSATGHT
 MDDFFELATLILIMLGHWIEMNAVGNAGDALKKMAELLPNNAIKVMDNGQREEVKISDI
 MTDDIVEVKAGESIPTDGIIVQGQTSIDESLVTGESKKVQKNQNDNVIGGSINGSGTIQV
 KVTAVGEDGYLSQVMGLVNQAQNDKSSAELLSDKVAGYLFYFAVSVGVISFIVWMLIQND
 35 VDFALERLVTVLVIACPHALGLAIPLVSTARSTSIGAHNGLIIKNRESVEIAQHIDYVMMD
 KTGTLTEGNFSVNHYESFKNDLSNDTILSLFASLESQSNHPLAISIVDFAKSKNVSFTNP
 QDVNNIPGVGLEGLIDNKTYKITNVSYLDKHLNYYDDDLFTKLAQQGNSISYLIEDQQVI
 GMIAQDQIKESSKQMIADLLSRNITPVMLTGDNNEVAHAVAKELGISDVHAQLMPEDEKE
 SIIKDYQSDGNKMMVVDGINDAPSLIRADIGIAIGAGTDVAVDSGDIIILVKNPSDIIH
 40 FLTSLNNTMRKMVQNLWWGAGYNIVAVPLAAGALAFI GLILSPAVGAILMSLSTVIVAIN
 AFTLKLK*

Sequence 2417

Contig_0754_pos_3900_3526,

45 putative peptide of unknown function

atggaaaatagcgtaagcaaagggttttttgggtgtagcaggttttgacttactagtagat
 gataataatgatgtttatgcgatttgatttaaaacttttaggcaaaacggatcaacgagtatg
 ctacttttagcaaaagatttaactcatggatatcataaattttacagttacttttctaatt
 ggagataatacaaaattctataatgctattttaaaatacgtagaattaggtgtactttat
 50 ccactttcctattacgatggagattggatggaaagaatcaagttaattctagatttggc
 tgcatttggcatggggaaaataaagaattaattaatcgatatgaacaacaatttatattg
 gaagctggattataa

Sequence 2418

55 MEIGVSKGFFGVAGFDLLVDDNNDVYAI DLNFRQNGSTSM LLLAKDLTHGYHKFYSYFSN
 GDNTKFYNAILKYVELGVLYPLSYDGDWYKQVNSRFGCIWHGENKELINRYEQQFIL
 EAGL*

Sequence 2419

Contig_0754_pos_3352_2687,
putative peptide of unknown function
atgtcatacaaatatgaagcatttttttaaagatatatttgattaatgaatatattttatttt
gcttcaaaaaataaaaaattaattagaatacaacatgagaatttgccatatattgctatg
5 tggacagacgaaaatgttgctgagctcttatttgttacatcattcaattgattacgacaaa
atcattagagcagatatgtgaccgttttgtaacatatgaaatggatgaaatctttgatcca
ggtgacaaagtttttagttaatgtgaataatggtgaagaaggaaacattgtagatatagtt
aaaatgactgatgagttgatgtctgaattagatgatataagaatgagagagtttatataaa
ga'gtcg:aaaaatatgacgaagtatacggattgacaaacaaaggtgaaaagaatt'catt
10 atgatttcagatgatgaccataacaaaccacacatcatgcctgtttggagtattaagagt
agagcgcgtaaaagtacgtgatcaagattttgaagaatgtgatttaatcgaaattgaaggt
gaagtccttagtgaatggtttagacaagttacgcgatgataataaagcagtagcgattgat
ttgaaatcaggtgttggtagtactgtgtatcagcgcaaaaactgtcaaatgaagcaaca
ttttaa

15 Sequence 2420
MSYKYEAFKIDILNEYIYFASKNKKLIRIQHENLPYIAMWTDENVAESYLLHHSIDYDK
IIRADIDRFVITYEMDEIFDPGDKVLNVNNGEENIVDIVKMTDELMSELDDIRMREFIK
DVAKYDEVYGLTNKGEKNFIMISDDDHKNKPHIMPVWSIKSRARKVRDQDFEEDLIEIEG
20 EVFSEWLDKLRDDNKAVAIDLKSGVVGTVVSAQKLSNEATF*

Sequence 2421
Contig_0754_pos_1366_569,
putative peptide of unknown function
25 atgtttataaaaaaggatttttgatgatattacagttcaagtatttgaagaaaaatataga
gatgcacttaaccaatttgaattaaagtgaacgacacaaatatattcttcattgcctcaa
actgttttagatgatgcattaaaagatgaaaatcgaattgctaattgtagctttaaataaa
gaaggaaaagtagtggggttcttcgttttgcatcggtattatcaacatgaaggttatgat
acacraaaacaatgttggtttatgtacgttcattgtcagttaatgaaaagtttcaag'ccat
30 ggatatgggacaaaaatgatgatgtttttaccagagtagtgcagcattatttctctgat
tttacacatttatacttagtagtagacgctgaaaaccaaagtgcttggaacgtttatgaa
cgtgcagggtttttatgcatacagctacaaaagaagaaggacctattgggaaagaagactt
tattattttagatttagattcaaaacatgtatcttctttaaggctaaaagagggggaagtc
acataataatgatgatattcacgtgattaatttgcttaagatgatgtaaaggtaggcttt
35 attgcactagaacaaaatgataataaaatgaatatttctgcaatcgaagttaataagaaa
aataggaatgaggggaattgcagaaaagtgctttacgccaattaccaacgtatatacgtaaa
cagtttgaagacattgaagttttatcaattactttagtgtccagcagcactaactcat
ttaatgtatctgcaataa

40 Sequence 2422
MFIKKDFDDITVQVFEEKYRDALNQFELSERQQIYSSLPQTVLDDALKDENRIANVALNK
EGKVVGFFVLHRYIYQHEGYDTPNNVYVRSLSVNEKFQGHGYGTKMMMFLEPYVQALFPD
FTHLYLVVDAENQSAWNVYERAGFMHTATKEEGPIGKERLYYLDLDSKHVSSLRLKEGEV
TYNDDIHVINLLKDDVKVGFIALEQNDNKMNISAEVNNKKNRNEGIAESALRQLPTYIRK
45 QFEDIEVLSITVLQHALTHLMYLQ*

Sequence 2423
Contig_0755_pos_1306_1917,
is similar to (with p-value 8.0e-42)
50 >gp:gp|AF012285|AF012285_40 Bacillus subtilis mobA-nprE gene
region. NID: g3282109. >gp:gp|Z99111|BSUB0008_137 Bacillus
subtilis complete genome (section 8 of 21): from 1394791 to
1603020. NID: g2633699.
atgacccaatataactttttcacctaaagatttttaaagcttttgaagtcgaaggttttagac
55 caaagaatggaagcacttaatgactatgtcagacctcaacttcatcaattaggatcttat
tttgaagaatatttctactacacaaacaggtgaaactttttatgctcacgtagctaaacac
gcacgtagaagtgtaatccacctatcgatacgtgggtagcttttgcctcctaataaacgt
ggttataaaaatgttaccacactttcaaactcgattgttttagaaatcagcttttcattatg
ttcgggtatcatgcacgaaggtagaataaagaagaaaagtgaaaatatttgataaacat

tttgataaactgacatctttaccaagtgattatagtgtttctctagatcatatgasaact
 gaaaagcactatatcaaggatatgagtaatgaagagttgcatgctgctatcgatagagtt
 aaaaatgttaaaaaaggtgaattttttgttgccagaacattatcaccaaccgataaaaaga
 ttaaaatctgataagtcttttctaaaatttgttgaggaaacttttgatgaatttttaaaa
 5 ttttatcaataa

Sequence 2424

MTQYTFSPKDFKAFEVEGLDQRMEALNDYVRPQLHQLGSYFEEYFTTQTGETFYAHVAKH
 ARRSVNPPIDTWVAFAPNKRGYKMLPHFQIGLFRNQLFIMFGIMHEGRNKEEKVKIFDKH
 10 FDKLTSLPSDYSVSLDHMKTEKHYIKOMSNEELHAAIDRVKNVKKGEFFVARTLSPTDKR
 LKSDKSFLKFVEETDFEFLKFYQ*

Sequence 2425

Contig_0755_pos_5604_4453,
 15 putative peptide of unknown function
 atgttaggagagcaatatacacaaattaagcgtccagcaaactcggttaactgaaaaaata
 ttagggttggttttagttgggtattcttactcatattaactattgtttcaatgtttattgcg
 ctctgtatcttttagtaaatgatacgtcaattgccaattagaaaacacacttaataataat
 gaactcgtacaacaaatttttagccaataatgatttaagtacaactcaatttgtgatttgg
 20 ttacaaaatggagtttgggcaattattgtttattttattgtttgtttgctcatctcgttt
 tttagcgttaatttctatgaatataagaattttgtctggtttacttttttaatagctgct
 atagtcacaattccgcttgtattgttgattgtaactctaattcctatcttattcttt
 atcattgcaatgatgatgtttgctagaagagatagaatagaaacagtgccatcttattat
 aatgaatatgatcaaccatactatgatgagagaggttttatgaaccagagtcagaaaat
 25 gaacatggatataatgatgatgtgtatgaacctatgcatactaaaaaggaagatagaaat
 acaagacgtcaattcaatagaaatgctcagcaacaagattcctataatgggtataactgat
 aatcaacccgatgaagatacatcttccgatcaactttattcagacgaatatgtagataat
 gaagataaatattctcaatttccaaaaagagcagttgaaagtgaatatgcatctcaacaa
 actgaagatgaaccaacagtcattgtcaagacaagctaagtacaataaaaaatctaaaaat
 30 acggattttgaagatgcgcaacaggaacatatggaaggtaatcaatttgatgacatagga
 gttgttgtaaccacaaattgatcttaagaactaaaaagcgcaagaaaaagagaaaaagca
 gaaatacgtgctaaagaaaaaagaaaaagcagataataaaacgtatgaaagaacga
 agaaaaaaccagccaagtgctgttaaccaacgacgtatgaattatgaagaacgtcgacaa
 atgattaataatgaacaagaagatacagataataacttaaatcaacaggaagattcaaaa
 35 aaagaaaattaa

Sequence 2426

MLGEQYTQIKRPANRLTEKILGWFSWVFLILITIVSMFIALVSFSNDTSIANLENTLNNN
 ELVQQILANNDLSTQFVIWLQNGVWAIIVYFIVCLLISFLALISMNIRILSGLLFLLIAA
 40 IVTIPLVLLIVTLIIPILFFIIAMMMFARRDRIETVPSYNEYDQPYDERGFYEYESRN
 EHGYNDDVYEPMHKTKEDRNTRRQFNRNAQQDSYNGITDNQPDEDTSSDQLYSDEYVDN
 EDKYSQFPKRAVESEYASQQTDEPTVMSRQAKYNKKSNTDFEDAQQEHMEGNQFDDIG
 VVEPQIDPKELKAQRKREKAEIRAKKKEKRKAYNKRKERRKNQPSAVNQRRMNYEERRQ
 MINNEQEDTDNNLNQQEDSKKEN*

45

Sequence 2427

Contig_0755_pos_1075_314,
 is similar to (with p-value 2.0e-65)
 >sp:sp|Q45499|SUHB_BACSU EXTRAGENIC SUPPRESSOR PROTEIN SUHB
 50 HOMOLOG. >gp:gp|AF012285|AF012285_41 Bacillus subtilis mobA-
 nprE gene region. NID: g3282109. >gp:gp|Z99111|BSUB0008_139
 Bacillus subtilis complete genome (section 8 of 21): from 13
 94791 to 1603020. NID: g2633699.
 atgatgcaagaagagtttagacattaaaactaaatcgaaaccaaattgatttagttacaaat
 55 gtggataaggcgacagagaattatctatatgaaacgattcttcataattatccagatcat
 caggttattgtcgaagagggacatgggtcataatctcgagtatttaagggggttatttgg
 gttattgatccaattgatggaacacttaattttgttcacaaaaagaaaaattttgccatc
 tctattggtatttatcatgatgggaagccttatgcagggtttgtttatgatgtcatgaaa
 gatgttttatatcatgcaaaggttggaacagggtgcatttgaaaatacacataaacttgaa

atgattcaaaaactgaacttaaaagaagtattataggtattaatcccaattggctgacg
aaaccaatactcagtgatatttttagttcaatagtgatgaggcaagaagtgcacgagca
tatggtagtgcagcattagaaattataagtgtagcgaaggggtcaattggcggcttaccta
acacctagactacaaccgtgggattttgcaggtggattggtgattttgaacgaagtaggt
5 gggataggaaccaacttattagcgataaattagacttcaatcaaccgaattcaatatta
atagcaaatcctagccttcctcgtgaaatattaaatcatcatttaaatcagcaagagat
acgcttattacacttcctgaaaaaagggtttgaaagagatag

Sequence 2428

10 MMQEELDIKTSNPNDLVTNVDKATENYLYETILHNYPDHQVIGEEGHGHNLEYLKGVW
VIDPIDGTILNFVHQKENFAISIGIYHDGKPYAGFVYDVMKDVLYHAKVGQAFENTHKL
MIQNTELKRSIIGINPNWLTKPILSDIFSSIVNEARSARAYGSAALEIISVAKGQLAAYL
TPRLQPWDFAAGLLILNEVGIGITNLLGDKLDFNQPSILIANPSLHREILNHLNQQRD
TLITLHEKRFGKR*

15

Sequence 2429

Contig_0756_pos_1371_1877,

putative peptide of unknown function

atgaaagacacaaacctaataattcgaaattaattcaaacatatttaagtaagaaaact
20 ttaagatatggtacagcaagtgcattaacattggcactctattttttaacagtaacgta
actgtgtatgcggatgaaaatactgcaaaccaaatcaaggaacatcaccaaaaacttca
cagacagcacctacaaataatactgaaaatacagatgccacagccataacaacagatcaa
aataataatgatgaagaagaatcagatgcgtcatatgaacttccaattctttatgtaact
gtctggctagatgatcaaggaatattattaaagatgctgtggaagatgctaaaaccct
25 gcttcagaaaggcaaccggtgaaaattcctgggtaccaacattatagaacttctgtgagt
gcggaattactaagtttattttatcgtaaaattagcactgcacaatcacctatagttgaa
aatcaaaaacgtatggtcgtggcatag

Sequence 2430

30 MKDNKPNNSKLIQTYLSKKTLYGTASALTALYLFNSNVTYADENTANQNQGTSPKTS
QTAPTNNTENTDATAITTDQNNDEEEYDASYELPILYVTVWLDQGNIIKDAVEDAKTP
ASERQPVKIPGYQHYRTSVSDGITKFIYRKISTAQSPIVENQKRMVVA*

Sequence 2431

35 Contig_0756_pos_4998_4510,

putative peptide of unknown function

atgccaaaagtacatcaagttaaggaaagatttgtgaaattaggggaccaacagtttaaa
gcatttgaaattagatacgtacatacattcattacgtgttgatgtgtgatggtgtagat
ttagcaatgaaacagcgctggaagattttgtcagtgcgcaaacatggcatcaacaattt
40 aaaacgattggcgtcatgctttttcaacaagataaacaattcatatatccactgatacat
atacctaaaatagatagcttaattctgggaaaatagctgtggttcaggagcggcttctatc
gggtgtgttagttaattatctaacagatcatgatattcaagattacctagttaaccaacc
ggaggcagattatttctcatccagaaagtctggacaaaatgaatacacaacagatt
aagtgtcaagtttcaactgtcgcaacaggacaagcatatatagaacaggagacaatgacg
45 caaatatga

Sequence 2432

MPKVHQVKERFVKLGDDQKFKAFFIYDYIHYVLMCDGVDLQKQVEDFVSAQTWHQQF
50 KTIGVMLFQDDKQFIYPLIHIPKIDSLIWENSCGSGAASIGVLVNYLTDHDIQDYLVNQ
GGSIIVSSRSKGQNEYQTTIKQVSTVATGQAYIEQETMTQI*

Sequence 2433

Contig_0756_pos_3690_2413,

putative peptide of unknown function

55 atggttggttagtgaccggtcgctattcaacttgctcgactatgtcatttacatggagaa
catatagttgatatggtgagtcggttcctgcatcaacaaatctaagagagtccttgat
gcttatcaacgtgacggcttttttcagtaatgactcaaaatgatgcacatcagtggttt
tcaggtaagtttacggttagacatttttttaagatgttaagatattactgaatattat
gacgtggtgatttttagcatgtactgccgatgcgtatcgaccgatattacagcaattatct

aagtcacattaaagcgtattaagcaaatacatcttgggtctaccaacattaggatcacat
 atgcttggttaagcaattactatcagatgttcaatgtgaaggtgaagtgtttcattttcc
 acttatctaggcgatacccgaaatatttgataaagcacaaccacattgtgtcctaa:caca
 cgagttaaatcaaaattatttcgtaggttcgactcaatctcagtcctatgacgttgtgtaag
 5 ctttaagtcctttatttgactatttgaatatagaattaacaacgatggacacaccactacat
 gggagatacataatagttcactttatgtacaccaccattgtttatgaatcaattttca
 ttaaaggcgttatttgaaggagcgaagtagtatatgtatataagctatttccagag
 ggtccaatcacaatgaccttaatacacgaaatgcgattaatgtggcaagaaatgatgatg
 atattaaaaaaattaaaggtaccttcggtcaatcttctaaagtttatggtgaaagaaaac
 10 taccctatacgttatgagaccatgcgcgaagtagatattgaaaactttaaaaatttaccga
 gctattcatcaagagtatctactttatgtgcgatatacagcaattttaatcgatccgttt
 tctaataccggacgatcaaggtgcatattttgatttttctgccgtaccatacaaacatggt
 gactatgatgaacaaggagtcatacatataccacgcagtcgcgagtgaagattattatcgt
 actttgataattcaagcatttgaagagcattaaacggttgcaacaccgatgattgacaca
 15 ttgttattacgttatgaaaatactgttaacaatactgtgacacacatttacatcaacaa
 ctatcaaggcaattcgaattacatcatttttaacaggatttagcggttagtgacgaactac
 ttaactttttataataaa

Sequence 2434

20 MVGSGPVAIQLARLCHLHGEHIVDMVSRVHASTKSKRVFDAYQRDGFSSVMTQNDHQC
 SGKFTVRHFFKDVKDITEYYDVVILACTADAYRPILQQLSKSTLKRIKQIILVSP
 TLGSH MLVKQLLSDVQCEGEVISFSTYLGDTRIFDKAQPHCVLTTRVKS
 KLFVSGTQSQSMTLCK LKSLFDYLNIELTMDTPLHAEIHNSSLYVHPPLFMNQFSLKAVFEGTKVPVYVYKLFPE
 GPITMTLIHEMRLMWQEMMMILKKLVPSVNLKFMVKENYPIRYETMREVDIENFKNLP
 25 AIHQEYLLVRYTAILIDPFSNPDDQGAYFDFSAPYKHVDTDEQGVIIHPRMPSE
 DYYR TLIIQAIGRALNVATPMIDTLLRLRYENTVKQYCDTHLHQQLSRQFELHHFKQDLALVTNY
 LTFYK*

Sequence 2435

30 Contig_0756_pos_2362_1841,
 is similar to (with p-value 3.0e-63)
 >gp:gp|AF076683|AF076683_1 Staphylococcus aureus oligopeptid
 e transporter putative substrate binding domain (opp-1A), ol
 igopeptide transporter putative membrane permease domain (op
 35 p-1B), oligopeptide transporter putative membrane permease d
 omain (opp-1C), oligopeptide transporter putative ATPase dom
 ain (opp-1D), and oligopeptide transporter putative ATPase d
 omain (opp-1F) genes, complete cds; and unknown gene. NID: g
 3800817.
 40 atgaataaactcacaaaactaagtcacagtcatttttgtatctggaattatttttagccggt
 tgtggaataacaaagaactaacagagaaaaagagaataaagtattatcatatacaact
 gtcaasagatattggagatatgaatccccatgtttatggaggttcaatgtcagcagaaagt
 atgatttatgagccgttagttcgcaataccaaggatggtattaagccattattagcaaaa
 aatgggacatttcacctgatggttaagacatatacgtttcatttaagggatgatgtatct
 45 tttcatgatggttacgaaatttgatgcagatgcagtgaaagaaaacatcgatgcagtacaa
 caaaataagaaactacattcatggttaagactttcaacactgattgatgatgtcaaagt
 aaggataagtatacagatacaactacatttgaagggaagcttatcaacctgcggttagcagaa
 cttagctatgccacgaccatacgtttttgattttcaactatag

Sequence 2436

50 MNKLTKLSTVIFVSGIILAGCGNNKELTEKKENKVLSTTVKDIGDMNPHVYGGSMSAES
 MIYEPLVRNTKDGIKPLLAKKWDISPDKTYTFHLRDDVSFHDGTFKFDADAVKKNIDAVQ
 QNKKLHSLRLSTLIDDVKVKDKYTIQLHLKEAYQPALAEALAMP RPYPVDFQL*

Sequence 2437

55 Contig_0757_pos_801_1223,
 putative peptide of unknown function
 atgagaaaatgggttaaccttactattaattacaacattgggtgttaactgcatgtggtaaa
 agtaacgaaaaagcttcttttagaaaaaagcattgatcagttgaaaaaagaaaataaggat

ttaaaaaacagaagaaaaagttacaagagcaaaaggataagcttaaacacaaacaggat
 agtctccaagaagatgtaaatgacttgctgctaaaagcacatcccgagataagaaaaat
 aaagataatcatgatgcaaaagaaaagcttcagataatcaatcgacatctgctaatacat
 gatgatcaaaactaacaaaaataaaaagcaatcaagatgaacatgacagtcaatcctctaaa
 5 ccacatacacagcagaagccctcacagaatgatagaaaaataatcatcgacaagaacga
 tag

Sequence 2438

MRKWLTLTLITLVLTAAGKSNEKASLEKSIDQLKKENKDLKKQKKKLQEOKDKLKHQD
 10 SLQEDVNDLPKSTSRDKKNKDNHDAKEKSSDNQSTSANHDDQTNKIKSNQDEHDSQSSK
 PHTQQKPSQNDKNNHRQER*

Sequence 2439

Contig_0757_pos_4581_4994,
 15 is similar to (with p-value 5.0e-18)
 >gp:gp|AF012906|AF012906_6 Bacillus subtilis yojP gene, part
 ial cds; yojQ/S, yojR, yojT, yojU, yojV, yojW, yojX, yojY, y
 oiz, and yokA genes, complete cds. NID: g2522404. >gp:gp|Z99
 114|BSUB0011_163 Bacillus subtilis complete genome (section
 20 11 of 21): from 2000171 to 2207900. NID: g2634230. >gp:gp|AF
 020713|AF020713_166 Bacteriophage SPBc2 complete genome. NID
 : g3025478.
 atgataggtacttatcaaagtataaaactttgaaatgatgaagacttttaagcattgg
 attcagactaatcattattggaatatgttgagaaatacgggtgtgttaggtatagcatta
 25 gataatcctctccacggttcaaagtaataatgatagatatgacgttgttttgagaatagat
 gaaacagtaaatgatcagacaatatctaaaagagattttacaggtggcatatatgtctgtg
 tttaaagtttagtcatacaaaaaataaatatagagaagttcttttagcaatttagaaaaatatt
 ttaaatgaaagtcatttgcgtatgagaaatgaaccaattatagagagatacattgaagaa
 gagggaaacagataaaagtgtgtgaaatgttagtgccatatctatgaagtaaattaa

Sequence 2440

MIGTYQSDKNFEMMKTFKHWIQTNHYWKYVEKYGLGIALDNPLHVQSNQCRYDVVLRID
 ETVNDQTISKRDFTGGIYAVFKVSHTKINIEKFFSNLENILNESHLMRNEPIIERVIEE
 35 EGTDFVCEMLVPIYEVN*

Sequence 2441

Contig_0757_pos_6065_5280,
 is similar to (with p-value 6.0e-63)
 >gp:gp|U87792|BSU87792_1 Bacillus subtilis tRNA-Ala, phospha
 40 tidylglycerophosphate synthase (pgsA) and ClnA (clnA) genes,
 complete cds, and RecA (recA) gene, partial cds. NID: gl842
 434.
 atgatacttgcgatgatatgtgggttaaagtcaactaatcttctcggttctcaatcagca
 ttcacatttaaagttgttatagcttaggttcagtatcttgcggccgctgggtatttaga
 45 gaacgcttcttagaaatcttatatttgcccaacataaacctgaaccttccacttcggga
 gacagacgttcaaaaccacgacgtctgaatttaatacatgtatttagtaggtatgggtccca
 gcagggttttaggtttttatttgatgatttaattgaaaaatacttatttagtgtagca
 acagcttctaattgggtttatttataggtgccatttatatgattatagctgataagattct
 aaaactgttcagcatcctcaaacagtagatcaaattaattatttccaagcatttgcatt
 50 ggtatctctcaagcaatagctatgtggcctggatttagtagatccggttcaacgatttca
 acaggtgttcttatgaaattgaatcataaagctgcatctgatttacttttattatgtcg
 gtaccaattatgttagctgcaagtggattatctttactaaaacattatgagtatattcat
 ttgacacataccattctacatttttaggatttttagcggcatttattgttggaatttaatt
 gcaatttaaacattcttacttaatttaataaaagtttaagtttagtaccttttgctattat
 55 agaattgtcttagttatttttatagcaatcctatacttcggattcgggtattggcaaagga
 atttaa

Sequence 2442

MILVDDMWLKSTNFLGSQSAFTFKVVIQLGSVFAAAQVFRERFLEILHIGQHKPEPSTSG

DRRSKPRRLNLIHVLVGMVPAGILGFLFDDLIEKYLFSVPTVLIGLFIGAIYMIADKYS
KTVQHPQTVDQINYFQAFVIGISQAIAMWPGFSRSGSTISTGVLMKLNHKAASDFTFIMS
VPIMLAASGLSLLKHYEYIHLAHPFYILGFLAAFIVGLIAIKTFLHLINKVKLVPAIY
RIVLVIFIAILYFGFGIGKGI*

5

Sequence 2443

Contig_0757_pos_3653_3204,

putative peptide of unknown function

atgaagactaacgtaattatagatggagatgcttgctcctgttgtaattctgtcattgaa
10 ttgacgaaagggacaggcattttttgtacaattttaagaagtttagccatttttcacaa
caaatacaacccgaacatgtaaaaattgtatacgttgatgacgggtcccgatgcggtagac
tataaaatagtcgaacttgctagcaataatgatatcgatcacacaagattatggactt
gctagtctactgatagacaaagtcatactgtcatgcataaaaggaaatattttatcac
tcaaacacatccaaagccttattaaatcaaagatatctaaatgctcaataagaagacga
15 ggtggtcgtcacaaaggccctcctcccttcacaacagaggatagacttaaattcgagcat
gcttttagaaaaatcattaatcaaatatag

Sequence 2444

MKTNVIIDGDACPVVNSVIELTKGTGIFVTILRSFHSFQQIQPEHVKIVYVDDGPDVA
20 YKIVELASNNDIVITQDYGLASLLIDKVHTVMHHKGNLYHSNNIQSLNQRYLNAQIRRR
GGRHKGPFPFTTEDRLKFEHAFRKIINQI*

Sequence 2445

Contig_0758_pos_2562_3488,

25 is similar to (with p-value 0.0e+00)

>gp:gp|AF072726|AF072726_1 Staphylococcus aureus putative he
me A synthase (ctaA) gene, complete cds. NID: g3320605.

atgatggggtgtttatcattgtttagaaagcaaaaccttaaatggttaggtgttttagct
acgattattatgacctttgtacaattaggtggcgccctcgtaactaaaacgggatcagaa
30 gatgggtgtggctcgtcttgcccttatgtaatggcgctttacttccagaaaatttacc
atacaacaattatagaactgagtcacgcgcagatcagccatttcacttatagttgta
ttatggcttgtaattacagcttggaacacattggatatattaaagaaatcaaaccactc
tctattattagtggtgtttttattagttcaagcacttgtaggtgctgctgctgtgata
tggaacacaaaatccttatgtattagcgctacattttgggtatttcacttatcagtttctc
35 tctgttttcttaattgacattaatttttctcaattgacaaaaaatatgaagctgacatt
ttattttattcacaaacctttacgtatcttaacttggttaattggctatcatcgatactta
actattttatacaggtgcttttagttagacataactaaatcaagtccttgcttatggtgcttg
cctattccatttgatgatatcgcttcctcataatgcgcgatgattgggtacaattttcgcac
agaggtatggcgctcatcacttttatctggattatgattacatttatacacgctattaag
40 aattattcagataatcgaaactgtacgttatgggtatactgcacatttatacttggtatc
cttcaagttattacaggtgctttatcagtcataactaatgtcaatttaattattgcgtta
ttccatgctttgtttatcacttacttattcggaatgattgcttattttattttactaatg
ttaagaacgacgagaagtcaaaaataa

45 Sequence 2446

MMGCLSLFRKQNLKWLGVLATIIMTFVQLGGALVTKTGSSEDGCGSSWPLCNGALLPENLP
IQTIIELSHRAVSAISLIVVLVITAWKNIGYIKEIKPLSIIISVGFLLVQALVGAAAVI
WQONPYVLALHFGISLISFSSVFLMTLIIFSIDKKYEADILFIHKPLRILTWLMAIIVYL
TIYTGALVRHTKSSLAYGAWPIPFDDIVPHNAHDWVQFSHRGMALITFIWIMITFIHAIK
50 NYSDNRTVRYGYTASFILVILQVITGALS VITNVNLI IALFHALFITLFGMIAYFILM
LRTRRSQK*

Sequence 2447

Contig_0758_pos_7092_3637,

55 is similar to (with p-value 0.0e+00)

>gp:gp|D83706|D83706_1 Bacillus stearothermophilus DNA for p
yruvate carboxylase, complete cds. NID: g1695685.

gtgtcttggttttgaacaaataaagaaattacttggtgctaaccgtggtgaaatcgcc
attagaatttttagagcggcagcagaattaaatcagtagcagtagcaattttatttcta

gaagataaaaagttcgttacatagatataaagcagatgaatcctatctagttggaagtgat
 ttaggacctgctgaaagttatltgaatatcgaacgtatcatcgaagtagctcttcgcgca
 ggtgtcgatgcaattcatcctgggtatgggttttttaagtgaataaagaacatttgcacgc
 cgatgtgctgaggaaggcattaaatttataggtccgcatcttgaaacatctagacatgttt
 5 ggagataaggttaaggctagaacaactgctattaacgctaacttacctgtaatccccgggt
 acagattggtcctattgaaagttttgaagctgcagaacagtttgctaataagcaggttac
 ccacttatgattaaagccacaacgcgtggcggttgtaaaaggtatgcgaatcgttcgtgaa
 tcaagcgaattagaagacgctttccatcggtgcgaatacagaagccgaaaagtcatttgg
 aatagcgaagtttatatcgaaagatatattgataatccaaagcatatagagggtcaagtt
 10 attggtgatgaattcggaatatcattcatttggatgaaagagattgctccgtacaacga
 cgtcatcaaaaggttgttgaaagttgcaccttcagtaggtctttctaacaataaagagag
 cgaatttggatgcccgaattcaactgatggaaaatataaaatacgtcaacgctggaaca
 gtgaatttttagtttctgggatgaatttttcttcattgaggttaattccacgtgttcaa
 gttgagcatcaataactgaaatgattactggtatagacattgtgaaaacgcaaatttta
 15 gttgctgatggagaatcgttatttggagataaaatctctatgccacagcaaaatgaaatt
 caaacattagggatgcatgataaatgctgtatatacaactgaagatcctactaatgtttt
 atgcaagattctggcacaattattgcatatcgatcaagtgccggttttgggtgtgaactt
 gatgcaggggatggattccaaggtgcagaaatttcaccttactacgattcactattagtt
 aagctttctacacatgccgtttcatttaacaagctgaagagaaaatggaacgttcatta
 20 cgcgaataatgcgaattcggtggcgtaaaagcgaatatccatttctcatcaatgttatgcgt
 atgtgataaatttagaagtggtgattataactactaaatttattgaaagaaacacctgaactt
 ttcgatattgcaccgacattggacagaggtaccaagacttttagagtatattggtaattgtg
 acgataaacggatttctaatgtagaaaagcgtccaaaaccagaatatgaatctaccaa
 atccccaaaatttctcaaaaagaaaatcaatcagttatttggaaacaaaacaaattcttgag
 25 caacatggaccaacaggtgttacaatttgggttagagaacaagaagatgttttaattacc
 gatactacatttagagatgcacaccaatctttacttgcaacacgtgtaagaacaaaagat
 atgatgaacattgcatctaaaactgctgaagtttttaaagatagtttttcattagaaatg
 tggggtggtgcaacatttgatgtcgctataatttcttgaaagagaatccatgggaacgt
 ttagaaagattgcgcaaaagccattccgaatgtgttattccaaatgttattacgagcttcg
 30 aacgcagtaggttataaaaactatcctgataatgtaattaagaaatcggttcagaaagt
 gcaaaagctgggtgtagatgttttccgtatattcgactcattgaaactgggttgatcaaatg
 aaagttagcgaatgaagctgttcaagaagctggaatggatctgagggtaacaatttgctat
 acaggtgatattttaaatgctgaacgttccaatatattatacttttagattattacgttaaa
 atggctaaaagaactggaaaagagaaggattccatatattagcaattaaagatatggctggt
 35 ttattgaaaccgaaagcagcttacgaatttaattgggtgaattacgtgaggcaacacatctt
 ccaattcatttacatcacatgatactagtggaaatggattgttgacatataaacaagca
 attgatgctggcgtagatattatagatactgctgttgcatctatgagtggtttaaagagt
 caaccaagtgcgaattcattatattatgcactaaatggatttccacgtaatttaagaact
 gatattgatgggttagaagagttgagtcattactggtctgtagtcagaccttactatgca
 40 gatattgagtgatatacaatcaccaatacagaaatttatcaacatgaaatgccaggt
 ggccaattatccaaacttaagtcaacaagctaaaagtttaggattgggcgaacgttttgat
 gaagtcaaaagagatgtatcgctggtgcaacttccgtgttgagatcttgtaaaagtaaca
 ccatcttcaaaggtagttggagatatggcactatatatgggtgcaaaatgatcttgatgaa
 gatacggtcatcaatgatggttataaattagatttcccagaatctgttggtgcatcttctt
 45 aaaggtgacattggacaacctgtcaacggattcaacaagaaatgcaagatgttatttta
 aaaggacagcaaccaattactgaaagaccaggtgaatacttgagccggtcgattttgaa
 gcaatccgtcaagaattaaagcgacatacaacaagacgaggttaacagaacaagatataatt
 agttatgtactttatccgaaggtatataaacaatatattcaaacgaaagagcaatttgggt
 aatgtatctttactggatacacccgacattcttatttggcatgcgtaatgggtgaaacagtt
 50 gaaattgaaattgatactggttaaagctctaattattaaattagaacaatcagtgaaacca
 gatgagaatggtaaacgtacaatttattacgctatgaatgggtcaagcaagacgtatttat
 attcaagatgaaaatgttaaaacgaatgctaattgttaaacctaagggcgataaatcaaat
 ccaaatcatattgggtgctcaaatgcctggttctgtaactgaagtcgaagtgctgttaggc
 gatgaagttcaagctaatcagccattattaatcactgaagcaatgaagatggaacgacg
 55 attcaggcaccatttgatggaattattaaacaataatgttgctaattggagatgccatt
 gccaaagagatttatttagtggaattgaaaagtaa

Sequence 2448

VSWLLKQIKLLVANRGEIAIRIFRAAAELNISTVAIYSNEDKSSLHRYKADES YLVGSD

LGPAESYLNIERIIEVALRAGVDAIHPGYGFLSENEQFARRCAEEGIKFIGPHLEHDMF
 GDKVKARTTAINANLPVIPGTDGPIESFEAAEQFANEAGYPLMIKATSGGGGKGMRIVRE
 SSELEDAFHRAKSEAEKSGFNSEVYIERIDNPKHIEVQVIGDEFGNIIHLYERDCSVQR
 RHQKVVEVAPSVGLSNKLRERICDAAIQLMENIKYVNAGTVEFLVSGDEFFFIENVPRVQ
 5 VEHTITEMITGIDIVKTQILVADGESLFGDKISMPQQNEIQTLYAIQCRITTEDPTNDF
 MPDSGTIIAYRSSGGFVRLDAGDGFQGAIEISPYDSLVLKSTHAVSFQAEKMERSL
 REMRIRGVKTNIPFLINVMRNDKFRSGDYTTKFIEETPELFDIAPTLDRGTKTLEYIGNV
 TINGFPNVEKRPKPEYESTKIPKISQKKINQLFGTKQILEQHGPVTNVWVREQEDVLIT
 OTTFRDAHQSLLATRVRTKDMNIAASKTAEVFKDSFSLEMWGGATFDVAYNFLKENPWER
 10 LERLRKAIPNVLFQMLLRASNAVGYKNYPDNVIKKFVHESAKAGVDVFRIFDSLNVVDQM
 KVANEAVQEAGMVSEGTICYTGDILNAERSNIYTLDYVYVMAKELEREGFHILAIDKDMAG
 LLKPKAAYELIGELREATHLPIHLHTDTSNGNLLTYKQAI DAGVDIIDTAVASMSGLTS
 QPSANSLYYALNGFPNRLRDTIDGLEELSHYWSVVRPYADFESDIKSPNTEIYQHEMPG
 GQYNSLSQQAQSLGLGERFDEVKEMYRRVNFVFGDLVKVTPSSKVVGD MALYMVQNDLDE
 15 DTVINDGYKLD FPESVVSFFKGDIGQPVNGFNKKLQDVILKGOQPITERPGEYLEPVDFE
 AIRQELSDIQQDEVTEQDIISYVLYPKVYKQYIQTKEQFNVSLDTPFLFGMRNGETV
 EIEIDTGKRLIIKLETISEPDENGKRTIYYAMNGQARRIYIQDENVKTNANVKPKADKSN
 PNHIGAQMPSGVTEVKVSVGDEVQANQPLITEAMKMETTIQAPFDGIIKQINVANGDAI
 ATGDLLVEIEK*

20 Sequence 2449
 Contig_0758_pos_2303_1368,
 is similar to (with p-value 3.0e-59)
 >sp:sp|P24009|COXX_BACSU PROBABLE CYTOCHROME C OXIDASE ASSEM
 25 BLY FACTOR. >gp:gp|Z98682|BS16823KB_3 Bacillus subtilis geno
 mic DNA 23.9kB fragment. NID: g2339988. >gp:gp|X54140|BSCTAB
 F_1 B. subtilis ctaB-F genes for cytochrome a assembly facto
 r and cytochrome-c oxidase (EC 1.9.3.1) subunits II, I, II,
 and IVB. NID: g994793. >gp:gp|Z99111|BSUB0008_160 Bacillus s
 30 ubtilis complete genome (section 8 of 21): from 1394791 to 1
 603020. NID: g2633699.
 atgagaaatttaaggagggaattatgaacaaagatcaaactttgtcacatactacgggc
 cgtgtatccttcaagaattacaacaaattattaaaatgggccttggtcaaggtaattta
 atacctgcttttgcaggcgcatggccttgcaatagtaatgacaaaccattccttctatct
 35 tccattccacaaatactattgatgctagttggctctacgcttattatggggggcgcttgt
 gctttaaataattattatgatcaagatattgatcgcatattgcctagtaagcaaagtaga
 ccaacagtaaatgatagaatatctgatagaaacttattaatgttaagttttgggatgatg
 ttaataggtgaagcatgtttattcttattaaatataccttctggtgttttaggattaatt
 ggtattgttggaatgtatcttactattcaatttggtctaagcgccatacaacttggaat
 40 actgttgttggaagtttctctggagctgtaccaccatttaattgggttggttagctatcgat
 ggatcattaagtttagcagcagtagcactcttttagttgtcttttgttggaacctatc
 catttctacgctctagcaattaaacgtagtgatgagtagcgcttgcaaatattcctatg
 ttaccatcagtgaaagggtttcaaacggacaagagtaagcatgtttatttggttagtgta
 ttattaccattgccattcttattatctaatttaggcgtaacttttgttggttattgctaca
 45 ctacttaatttaggatggtagcttttaggtttacaacggttcagaaaagaatctaataca
 actaaatgggcaacgcaaagtgttcgtttattcattgaactacttagtagtattccttgca
 ctrgttgtagttgtttcattaatcaagatgatataa

Sequence 2450
 50 MRNLRRGIMNKDQTLSTTTGRVSFKELQQI IKMGLVQGNLIPAFAGAWLAIVMTNHSFLS
 SIPQILLMLVGSTLIMGGACALNNYYDQIDIRIMPSKQSRPTVNDRISDRNLLMLSFGMM
 LIGEACLFLLNIPSGVLGLIGIVGYVSYSIWSKRHTTWNTVVGSFPGAVPPLIGWVAID
 GSLSLAAVALFLVFCWQPIHFYALAIKRSDEYALANIPMLPSVKGFKRTRVSMFIWLVL
 LLPLPFLLSNLGTFVVIATLLNLGLWALGFTTFRKESNQTKWATQMFVYSLNYLVVFFA
 55 LVVVVSLIKMI*

Sequence 2451
 Contig_0758_pos_1351_881,
 putative peptide of unknown function

atgaaacttatgaacgtacccatTTTTaccaacgataagtacatcgtgtattgttattagt
gcgatttttagtcgctattgggtgggcattgatttggaaacgtcaagttcataagcataaa
aacattatgctatgggctgcctttttcgcctttaacattctttattattatgcagcaaga
actatttttatcggttaatacagcttttcggtggaccaagttctattaaagtttattacact
5 attttcttagttttccatatcattcttgcacacagttgggtggcgttttaggcttaattcaa
atcatttttagccttcaaagataaaacttcatattcacagaaaaattgggccttgggcttca
ataatttgggttctttaccgcaattactgggtgttgcagtttatgtattgttatatgtattg
tatccaggtggagaaacaacatcattgcttaaagctacattaggtctataa

10 Sequence 2452

MKL MNVPILPTISTSCIVISAILVAIGWALIWK RQVHKHKNIMLWAAFFALTFFIIYAAR
TIFIGNTAFGGPSSIKVYYTIFLVFHIILATVGGVLGLIQIILAFKDKLHHRKIGPWAS
IIWFFTAITGVAVYVLLYVLYPGGETTSLKATLGL*

15 Sequence 2453

Contig_0759_pos_5038_3959,

is similar to (with p-value 2.0e-29)

>sp:sp|P33642|YFIT_PSEAE_HYPOTHETICAL 39.5 KD OXIDOREDUCTASE
IN FIMT 3'REGION (DADA*) (ORF2). >gp:gp|L48934|PSEPILRV_2 P
20 pseudomonas aeruginosa (isolate pRIC351) pilR gene, 3' end of
cds, dada*, fimT, fimU and pilV genes, complete cds. NID: g
1161217.

atgtcaattgctagacacctcagtgcaacacacttagatggttgcagtcatagatagagat
gtacctggaaagcatgcgtcatataaagctggaggtatgcttggcgacacaaatgaattt
25 acagaggtatgtagcttgtttcaattagccatcgaatctcgtgctatgtttccacaatta
agtaaatcattattagatgaaacagggcatagacattcaatttaaaaaattcaggacttatc
aaaattgctaataaacacgatgatattctcatctataaaacgacaatatcaatttctgaat
agtcaagaccgttagtgcaacaattatcagatgatgatttgctacaacttacacatggt
gaagttaaaccttcatacgcggccattcacataccacacgatgggtcaaattaatgcacat
30 cattacacactggcattattagaatcaatgaagttaagagatattaagcgttatgagtct
acagaggtcacttcaatagaacggcataatggctattattcagtgaaaaaccgatcaatct
tcaacaattgaaagcgacacaaaattatcgttgcaggtggcgcatggtcttcgcaattatta
acacaatatcatctacaacgacaagtgattggcggttaaagggtgaagttatcttattagaa
aataacgatctttcacttactgagacattatttatgactaatgggtgttacatcggtcca
35 aaacaacccaatcgttttttaattgggtgcgacgagtgaaatttaataattattctgtcggt
actacagatgaaggtatggattggcttcttcgccatgcatatcatcgtgtacctcaacta
aaagacagtcataactgaagaaatggtcaggagtaagaccatacacagaaaaagaaatg
ccagtcagtgatcaaattgatgaggttatacgtgataagtggtcattatcgaaaacgga
atatatgtcacctattatcggtcgtgacattgccaatgggtacttctgggtatataa
40 ccatcacgttattcaagttttacagttacaaggaggaataatcatgaagtgtatcattaa

Sequence 2454

MSIARHLSATHLDVAVIDRDVPGKHSYKAGGMLGAQNEFTEDSDLFQLAIESRAMFPQL
45 SKSLLDETGIDIQFKNSGLIKIANEHDDISSIKRQYQFLNSQDRSVKQLSDDDLLQLTHG
EVKPSYAAIHIPHDGQINAHHYTLALLESMKLRDIKRYESTEVTSIERHNGYYSVKTDQS
STIEAHKIIIVAGGAWSSQLLTQYHLQRQVIGVKGEVILLENNDLSLTETLFMTNGCYIVP
KQPNRFLIGATSEFNYSVGTDEGMDWLLRHAYHRVPQLKDSHILKKWSGVRPYTEKEM
PVMDQIDDGLYVISGHYRNGILLSPIIGRDIANWLLSGIKPSRYSSFTVTRRNHEVYH*

50

Sequence 2455

Contig_0759_pos_3004_2006,

is similar to (with p-value 2.0e-52)

55 >gp:gp|AF012285|AF012285_3 Bacillus subtilis mobA-nprE gene
region. NID: g3282109. >gp:gp|Z99111|BSUB0008_99 Bacillus su
btilis complete genome (section 8 of 21): from 1394791 to 16
03020. NID: g2633699.
atgtcgcgttatgaacgtcaaacacgctttgcaccatttggagaagagggtcagc:aaag

ctatcctcctctcaaaacttatttttggcgctggtgcttttaggaagccatattgtagat
 caactcgacgcatgggggctcatcatattgcaatcgctcgatatggatattggtgaaatt
 tcaaatttacatcgacaaacactcttcgatgaagaagacgcacatactttaatatccaaa
 gttgaagcaatcaagcataaggttaatcaaattaataaaatgtcaatctaacaacttat
 5 gatttagaagttacttcatcaaatatcgaaaatttgataaaaaatgtcgaaccagacatc
 atcattgatggcatggataaacttcaaaatcacgataacctgattaatgaggtttgtcacaag
 tatcaaatcccatgggtttatggtgcagctgttggtagtaaaggatcagtatatggaata
 gatcaccaaggaccatgtctaaaatgtttattgcaaacaattcctgacacaggggaaagt
 tgcgctattaatggcgtaattccccctgttatatcaatgattgcaagctatgaagtagca
 10 gaggccgtacgttatctttcaggaaaaggattttcaaagcaattaatcactattgaagca
 tttaatatcaattataagtcaatgaatgtagatgcactcaaaaataaagattgcccagtg
 tgtgaaaaacatgaatatagttactagaaagccaacaagaacgtactattgaggacttg
 tgtgggaatgctttatttttagattccccacctaagcttttaaacacgctgcccatttc
 cctgggaatatggtgaaatctacttcctttgccaaattaattcaatatcaaaccttatgaa
 15 ttcaccttgtttaagatggtcgatgaatgcatatggtatacacaatgatgaagaagca
 catcacctatacaatacgttggttaaaatccatacgttaa

Sequence 2456

MSRYERQTRFAPFGEEGQQLSSSQILIFGAGALGSHIVDQLARMGAHHIAIVDMDIVEI
 20 SNLHRQTLFDEEDAHTLISKVEAIKHKVNQININVNLTYYDLEVTSSNIENLIKNEPDI
 IIDGMDNFKIRYLINEVCHKYQIPWVYGAAVSGSKGSVYGIDHQGPCLKCLLQTIPTGES
 CAINGVIPPVISMIA SYEVAEAVRYLSGKGFSLITIDAFNINYSKMNVDALKNKDCPV
 CEKHEYTLLESQQUERTIEDLCGNAYLFRFPKAFKHAHFPGNMVKSTSFAKLIQYQTYE
 25 FTLFKDGRMNAYGIHNDEEAHHLNTLLKSIR*

Sequence 2457

Contig_0759_pos_1898_1284,
 is similar to (with p-value 4.0e-38)
 >sp:sp|Q48630|APL_LACLA ALKALINE PHOSPHATASE LIKE PROTEIN. >
 30 pir:pir|S39339|S39339 alkaline phosphatase-like protein - La
 ctococcus lactis >gp:gp|Z29065|LLALPHLP_2 L.lactis (MG1363)
 apl gene for alkaline phosphatase like protein. NID: g435295

atggaacaaattatcactgattttattagtaagtgggtttatacagcgatattcatttta
 35 atcttattagagaacgtattacctgtcgttccatctgagattattttaacttttgcaggc
 ttattatctgtgaaatcacacttatctatttggacattattaatcatagcaacaattgct
 tcattcattgggtttactcattttgtattatattttagacttatctcagaagagaaatta
 tatcgtttcgttgatcgacatggttaagtggatgaagttaaaaaagtaaagatttgaaacgg
 gcaaatgattggtttaaaaagtatggtgcgtgggctgtattttatgtcgttttgcacca
 40 gtaacttcgagtattaattacaataacctgctggcattaatcgaatgaacgttatcacgttt
 acaactttatctttaataggtactacaatttggaattttgctttaatactgctcggtcgt
 ttgctcagtgacagttttgacgctttgatgaatggtattcatatattcacgtatcatg
 tatgtcattattattattgcagtcataatatttgttatacgttatttaataagaaacgtcgt
 45 cggagtgttaataa

Sequence 2458

MEQIITDFISKWGYTAIFILILENVLPVVPSEIILTFAGLLSVKSHLSIWTLIIATIA
 SFIGLLILYYICRLISEEKLYRFVDRHGKWMKLKSKDLKRANDWFKKYGAWAVFLCRFVP
 50 VLRVLITIPAGINRMNVIQFTTSLIGTTIWNFALILLGRLLSDSFDALMNGIHTYSRIM
 YV:IIIA'YFVIRYLMKRRRSVK*

Sequence 2459

Contig_0759_pos_1086_151,
 is similar to (with p-value 0.0e+00)
 55 >gp:gp|D78193|BACGNTZA_11 Bacillus subtilis 36kb sequence be
 tween gntZ and trnY genes encoding 34 ORFs. NID: g1064780. >
 gp:gp|Z99124|BSUB0021_124 Bacillus subtilis complete genome
 (section 21 of 21): from 3999281 to 4214814. NID: g2636442.
 atgaaacaaaaatatctagatttactctcacaaaaatttgacagtgacagaaaaacttgct

actgaaattattaacttagagtcgaatcttagaattacctaagggaactgaacattttgtt
 ag'.gaccctcatggggaatcacgaatctttccaacatgttttaagaaacggatctg'aaat
 gtgcgtgctaaaaattaatgatctctcaaagataaattatcccagcaagaaatcaacgac
 ttagcagcattagtagtatactatccggaagaaaaactaaaattagttaaaaataatttcgat
 5 tcaatcggaacattaaatatttgggtatattacaaccattcaacgattaattgatttaatt
 acatattgctcatcaaaaatacacagttcaaaattacgcaaagcattacctgaacaatac
 gtttatattattgaagagctactttacaagagcaatgaatttcataataaaaagccttat
 tatgaaacattagtttaaccaaattattgaattagaacaatcagatgatttaattcattggc
 ctttcctatactgtacaacgtctagtcgttagaccatcttcagtcgtggcgatctctat
 10 gaccgtggctcctaaacctgataagattatggatacattaataaattatcattctgtagat
 atccaatggggaaatcatgatgtattatggattggcgctatgctggttcaaaagtatgt
 ctgctaaccttctacgtatctgtgcacgttatgataatttagatattattgaagatgca
 tatggcatcaatctacgccccttacttacgcttgctgaaaagtattacgatgctgaaaac
 ccagcggttaaacctaagaaacgaccagataaagacgtcagtccttcaaaaacgcgagaaa
 15 gtcaaatcacaaaaattcatcaagcaattgcatga

Sequence 2460
 MKQKYL DLLS QK FDS AEK LATE I INLES ILEL PKG TEH FVSDLHGEYESFQHVLNRNGSGN
 VRKINDIFKDKLSQEQEINDLAALVYYP EEEKLKLKNNFDSIGTLNIWIYITTIQRLIDLI
 20 TYCSSKYTRSKLRKALPEQYVYIIEELLYKSNEFHNNKPPYYETLVNQIIELEQSDDLIIG
 LSYTVQRLVVDHLHVVDIYDRGPKPKIMDTLINYHSVDIQWGNHVDLVIGAYAGSKVC
 LANLLRRCARYDNLDIIEDAYGINLRPLLTAEKYDAENPAFKPKKRPDKDVS LTKREK
 VKSQKFIKQLR*

25 Sequence 2461
 Contig_0762_pos_5747_4368,
 is similar to (with p-value 0.0e+00)
 >sp:sp|P25811|THDF_BACSU POSSIBLE THIOPHENE AND FURAN OXIDAT
 ION PROTEIN THDF. >pir:pir|JQ1215|JQ1215 hypothetical 50K pr
 30 otein - Bacillus subtilis >gp:gp|D26185|BAC180K_60 B. subtil
 is DNA, 180 kilobase region of replication origin. NID: g467
 326. >gp:gp|X62539|BSORIGS_5 B.subtilis genes rpmH, rnpA, 50
 kd, gidA and gidB. NID: g40020. >gp:gp|Z99124|BSUB0021_207 B
 acillus subtilis complete genome (section 21 of 21): from 39
 35 99281 to 4214814. NID: g2636442.
 atggattttgatacagattacaagtatttcaacaccgatgggtgaagggtgctattggaatt
 gtgagattatctgggccacaagctattgaaatcgagatatcttatataaaggtaagaaa
 aagttatctgaagttgagacgcatacaataaattacggtcatattattgatccagaaaca
 aatgaaacagttgaagaagtcattggtgtctgtattacgtgcccctaaaactttcacacga
 40 gaagatattattgagataaattgtcattggtggtattttaaacaattaatcgatatattagag
 ttactatgacttatggtgcacgtatggcagaaccaggtgaatatacaaaacgtgcattt
 ttaaatggtcgatatagattttatctcaagcagaagcggttatggattttatacgtt:caaa
 actgatcgagcttctaaggttgcgatgaatcaaatagaaggacgtttaagtgacttaatc
 aagaaacaacgtcaatccatattagagatactcgccaagttgaagttaacattgattat
 45 ccagagtatgatgatgtagaagacgcaacgacggacttcttactagaacagtcataagcgt
 attaaagaagaaatcaatcagttacttgaaacaggagcacaaggtaaaataatgagagaa
 gggttatctacagttattgtaggacgtcctaattgttgggaagtcctcgatgctaaataac
 cttattcaagataataaagcaattgtgactgaggtcgctggtacaacaagagacgtgtta
 gaagaatatgtcaatgttagaggtgtcccgttacgactttagatactgcgggtattagg
 50 gatactgaagatatcgtagagaagattggtgtagaacgttctaggaaagctttaagtga
 gcagatttaattttatttgtgcttaataacaatgaacctctgacggaagatgatcaaact
 ttattcgaagtcattaaaaatgaggatgttattgtaatcattataaaacagatttagaa
 cagcgatttagatgttagcgaactaagagagatgattggtgatatgccacttatacaaca
 tcgatgcttaaacagaaggtattgatgaattagaatacaaaattaaagatttattcttt
 55 ggtggcgaagtacaaaatcaagatatgacttatgtatctaattcacgtcacatttcattg
 ttgaaacaagcgagacaatcaattcaagatgcgattgatgctgctgagtcgtggtatccca
 atggatatggtacagattgatttaacacgtacttgggaaattctaggagaaattattgga
 gaatcagcgagtgatgaattaatagatcaactatttagtcaattttgttaggaaaataa

Sequence 2462

MDFDTITSISTPMGEAIGIVRLSGPQAIEIGDILYKGGKKLSEVETHYTINYGHIIDPET
 NETVEEVMVSVLRAPKFTFTREDIIEINCHGGILTINRILELMTYGARMAEPGEYTKRAF
 5 LNGRIDLSQAEAVMDFIRSKTDRASKVAMNQIEGRSLDIKKQRQSILEILAQVEVNIDY
 PEYDDVEDATDFLLEQSKRIKEEINQLLETGAQKIMREGLSTVIVGRPNVGKSSMLNN
 LIQDNKAIYTEVAGTTDRDVL EYVNVVRGVPLRLVDTAGIRDTEIVEKIGVERSRKALSE
 ADLILFVLNNNEPLTEDDQTLFEVIKNEVDVIVIINKTDLEQRLDVSELREMIGDMPLIQT
 10 SMLKQEGIDELEIQIKDLFFGGEVQNQDMTYVNSNRHISLLKQARQSIQDAIDAESGIP
 MDMVQIDLTRTWEILGEIIGESASDELIDQLFSQFCLGK*

Sequence 2463

Cortig_0762_pos_4295_2418,
 is similar to (with p-value 0.0e+00)
 15 >sp:sp|P25812|GIDA_BACSU GLUCOSE INHIBITED DIVISION PROTEIN
 A. >pir:pir|JQ1216|BWBSGA gidA protein - Bacillus subtilis >
 gp:gp|D26185|BAC180K_59 B. subtilis DNA, 180 kilobase region
 of replication origin. NID: g467326. >gp:gp|X62539|BSORIGS_
 6 B.subtilis genes rpmH, rnpA, 50kd, gidA and gidB. NID: g40
 20 020. >gp:gp|Z99124|BSUB0021_206 Bacillus subtilis complete g
 enome (section 21 of 21): from 3999281 to 4214814. NID: g263
 6442.
 gtgggttcaagaatatgatgtagtagtcattgggtgctggtcacgccggtattgaagcaggt
 ctagcttcagctcgccgggtgctaaaacactgatgtaacaattaatttagataatatt
 25 gctttcatgccatgtaatccatctgtaggtggtcctgcgaaaggatcggtgtacgtgaa
 atagacgcttttaggtggacaaatggcaaaaactattgataaaactcacattcaaatgcgt
 atgcttaatacaggtaaagggtccagctgttagagctttacgtgctcaagcagataaagta
 ttatatcaacaagaatgaagcgtgtacttgagaatgaggataatttagacatcatgcaa
 ggtatggttgatgaactcattatagaagataatgaagttaaagggttcgtactaatatt
 30 ggtacagaatatcggttctaaagctgtcattattacaacaggtacattcttacgtggagaa
 attatactaggaaacttaaaatattctagtggccctaaccatcaattaccatctgtaact
 ctagecgataatttaagaaaattaggatttgatcggttagatttaaaacgggtacacca
 ccacgtgtaaatgcgagaaccatcgattattctaaaactgaaatccaaccaggtggtgat
 ataggtcgagcggttagttttgaaacaaccgaatttatttttagatcaattaccttgttg
 35 ttaactttacaagaatggagatacacatcaagtcattgatgataacttacatttatctgct
 atgtattccggtatgattaaagggtacaggtcctagatattgtccatcaattgaggataaa
 tttgtccgctttaacgataaaaccaagacatcaacttttcttagaacctgaaggacgta
 acgaatgaggtatagctgcaaggattatctactagttacctgaacatggtcaacgtcaa
 atggttagaaaccattccaggtccttgaaaaagcagatatgatgcgtgcgggttatgctatt
 40 gaatatgatgcaatcgctgctactcaattatggccaacggttagaaacaaaagcgattaa
 aacttgatactcaggtcagattaatggaacatcaggatatgaagaagcagcgggacaa
 ggaatcatggcaggtattaaacgctgctggtaatgttttaggtacaggtgaaaaaatactc
 agccgttcagacgcataatattgggtgtacttatagatgatttagtcactaagggtacaaat
 gaaccgtatcgattattaacttcacgtgcggaatcgattattactacgtcatgataat
 45 gctgatttactgcttactgatattgggttatgaattagggttaatatcagaagaacgctat
 gcaagatttaataaaaagcgtcaacaaatcaaagatgaaatacaacgacttaccgatgta
 cgtattaaacaaatgaacatacgaagcaattattgaagctaagggtggttcaagatta
 aaagatggcatattagcgattgatttattacgtcgccccgaaatgaactacgaaacaatt
 ttagaaatcttagaagaatcacatcaacttcctgaagcgggttaggaacaagttgaaatt
 50 caaacaataatgaagggttatatcaataaatctttacaacaagtagaaaaagttaaaaga
 atggaagcgaaaaaattcctgaggatttagattatagcaaggtagatagtttagcatct
 gaagcagcgcaaaagttagctgaagttaaaccattaaatattgcacaggttcacgaatt
 tcaggtcggaatccagcagatatctcaattctactgtttatttagaacaaggtaaacct
 55 caaagggtgaaacaataa

Sequence 2464

VVQEYDVVVIGAGHAGIEAGLASARRGAKTLMILTINLDNIAFMPCNPSVGGPAKGIVVRE
 IDALGGQMAKTIDKTHIQMRMLNTGKGPAVRALRAQADKVLVYQQEMKRVLENEEDNLDIMQ
 GMVDELIIEDNEVKGVRTNIGTEYRSKAVIITGTGFLRGEIILGNLKYSSGPNHQLPSVT

LADNLRKLGFDIVRFKTGTPPRVNARTIDYSKTEIQPGDDIGRAFSFETTEFILDQLPCW
 LTYTNGDTHQVIDDNLHLSAMYSGLKGTGPRYCPSIEDKFVRFNDKPRHQLFLEPEGRN
 TNEVYVQGLSTSLPEHVQRQMLETIPGLEKADMMRAGYAIEYDAIVPTQLWPTLETKAIK
 NLYTAGQINGTSGYEEAAGQGIMAGINAAGNVLTGTEKILSRSDAYIGVLIDDLVTKGTN
 5 EPYRLLTSAEYRLLLRHDNADLRRLTDMGYELGLISEERYARFNEKRQQIKDEIQRLTDV
 RIKPNEHTQAIIEAKGGSRLKDGILAILDLRRPEMNYETILEILEESHQLPEAVEEQVEI
 QTKYEGYINKSLQQVEKVRMEAKKIPEDLDYSKVDLSASEAREKLAEVKPLNIAQASRI
 SGVNPADISILLVYLEQGKLRVKQ*

10 Sequence 2465

Contig_0762_pos_1424_816,
 is similar to (with p-value 1.0e-40)
 >sp:sp|P37524|YYAA_BACSU_HYPOTHETICAL_32.8_KD_PROTEIN_IN_SPO
 0J-GIDB_INTERGENIC_REGION. >pir:pir|S18078|S18078_hypothetic
 15 al_protein_3 - Bacillus subtilis >gp:gp|D26185|BAC180K_57_B.
 subtilis DNA, 180 kilobase region of replication origin. NI
 D: g467326. >gp:gp|X62539|BSORIGS_8_B.subtilis_genes_rpmH, r
 npA, 50kd, gidA and gidB. NID: g40020. >gp:gp|Z99124|BSUB002
 1_204_Bacillus_subtilis_complete_genome_(section_21_of_21):
 20 from 3999281 to 4214814. NID: g2636442.
 atgttcgaaattatagccggcggaacgacgatttcgagcattacagtcggttcataaacct
 caagtagatgtcattgttcgagatattgagatgaagaaacagcggtagttgcattgatt
 gaaaattattcaactgtaaaacttattctgtcgtcgaagaagcggaagcttataaaaagtta
 cttgaaatcgggggaacgactcaaaatgaattagcaaaaagcttaggcaagagccaagc
 25 ttcatgtctaataaacttagattattgaagtttagcacccaatgtgattaagagatracgt
 gaaggttaagattacagaaagacatgcacgagcggtattagttacctgatgaaacacaa
 gaagaattaatcgagcaagttatttagtcagaagttgaatgtgaaacaaactgagtaga
 gtacgtcagaaaaactggaccagagaaaagtaaaagcgacgactttccaattttctcaagat
 gtaacacaagcaaaaagaagaactaggttaagagtattgaaacgatagaaaaatcaggtata
 30 cgcgttgaacaaaaagataaagaacatgaagattattatgaaattaaaataaagatatat
 aagaataaa

Sequence 2466

MFEIIAGERRFRALQSLHKPQVDVIVRDMDEETAVVALIENIQRENLSVVEEAEAYKKL
 35 LEIGGTTQNELAKSLGKSQSFIANKLRLLKLAPNVIKRLREGKITERHARAVLVLPDETQ
 EELIEQVISQKLNVKQTEDRVRQKTGPEKVKAQTFQFSQDVTQAKEELGKSIETIEKSGI
 RVEQKQKEHEDYYEIKIKIYKK*

Sequence 2467

40 Contig_0763_pos_624_1304,
 is similar to (with p-value 2.0e-26)
 >gp:gp|U76260|PAU76260_1_Peptostreptococcus_asaccharolyticus
 alpha- and beta-subunits of L-serine dehydratase (sdhB) and
 (sdhA) genes, complete cds. NID: g2315864.
 45 atggtcaaaaagctatgattatcaaagtgccttcgatattattggaccagtaatgatggga
 ccttcaagttctcatcacagcaggtgcagtaaaaattgtaattcagcgagagctgtgtta
 ggagatattgcctaagcatatagaaattcgttattatgaatcttttgctaaaacgcatcaa
 gggcatggtagacagacgttgctattgtcggaggtgctatgggctacagcactttcgatagt
 agaattaaatcatccttagacatagcaaaaagatgaaaatattacaattgatattattgaa
 50 gatgaaggagaaaagtattggtcaacatcctaactgtgcttatatcaaagcaaaatcgaaa
 gacggacgttatatagaagtgataggtatttctattggtggcggtacaatcaaactaaaa
 ggtatcaatgtaaatggtttaaatgtggaactgaatcatgggcttccaatggttagttata
 gatggaaatattgaataaagctaaaaataaatcatcttattaatgatttatcagatatggac
 ttagacttaggtgaagaattaatagaacaaatgataatgaaggtttagttgtatttcct
 55 ttaataaagcaatctcagaatcagcattaaatattatttaaagataaacatagtgattta
 aacgtttcctatatcaaatag

Sequence 2468

MAKSYDYQSAFDIIGPVMMGPSSSHTAGAVKIGNSARAVLGDMPKHIEIRYYESFAKTHQ

GHGTDVAIVGGAMGYSTFDSRIKSSLDIAKDENITIDIIEDEGESIGQHPNCAYIKANTK
 DGRYIEVIGISIGGGTIKLGKINVNGLNVELNHGLPMLVIDGNMNAKINHLINDLSDMD
 LDLGEELIETNDNEGLVVFPLNKAISESALNIIKDKHSDLNVSIIK*

5 Sequence 2469

Contig_0763_pos_3838_4449,

putative peptide of unknown function

atgaagcagtggaatgaatagattaatcaccttaataggcgtattgttaatcatttttagct
 atttatttattctcaaagccatatatcgataattatctacatgaaaaagataaacgatcat
 10 aaaattgaaaattatgataaaaaaggaaaaagaacagacaaagacatctaaatcgacgcca
 aagataccttccgataaaatctaaaatggctggttatatagaagttccagatgcacaaata
 aaagaaccagtataccctgggtccagcaacaccagaacaactcaatagaggtgttagtttt
 gcagaaggtgacgaatctcttaatcaacagaatatttcaattgctgggtcatcggtttaca
 gatcgttcgcactatcaattttacaaattttaaaatcagccaaaatcggtagttaaagtgtat
 15 tttaaaactggaaatcaactagaaagtataaaataactaaaatacgtgatgttaagcct
 acagaggttaaggtattagacgaacatcctaataagaaaaatcaattaacattaaactact
 tgcgatgactataacgaagaaacgggtgtttgggaaacaaggaaaatattcatagctaca
 caaattaactaa

20 Sequence 2470

MKQWMNRLITLIGVLLIILAIYLFSPYIDNYLHEKDNHDKIENYDKKEKEQTKTSKSTP
 KIPSDKSKMAGYIEVPDAQIKEPVYPGPATPEQLNRGVSFAGDESLNQONISIIAGHTFT
 DRSHYQFTNLKSAKIGSKVYFKTGNQTRKYKITKIRDVKPTEVKVLDEHPNKKNLTLIT
 CDDYNEETGVWETRKIFIATQIN*

25 Sequence 2471

Contig_0763_pos_4794_5603,

putative peptide of unknown function

atgattaaagccattgcggtagatatggatggaacatttcttgacacaaataaacagttt
 30 gatcgaaatcggttttgaaactatttttaagaattaatagataaaaaatattaagtttata
 gctgcgagtggaatcaatttgcaaagctaaaatcaatttttgagatagggaatgttc
 tttatatctgaaaatggagcagtcattctataaaaggaatcaactttacaattatcgaagt
 tttgatcagtatatttttcaaaaagttgtaaattattttaaatttgatcaaaaagataaac
 aatttgattatttgggtgtaaaaagtgcatatattttaaaagaaacaagcgaagcattt
 35 aagcaagatgcacgtacatattatcaccaactaatagaggttgactccttacaacatta
 cctgatgatgattatgtgaaaattgctttcaatataaatcgtcagactcatccagactta
 gatgagaaattagctcttaagtttaaaagcagatattaaactagtatcaagtgggagagat
 agtatagatgttattatgccaaatgatgactaagggtcaagctttgtctagattattaaaa
 gaattggcaaatcctgcttcacatttaattggcatttgagatgcaaaataacgataaagat
 40 atgttgagagcttgccgaacatagttatgttatggctaatagtgaagatcaatcattattt
 aatatagcaggtcatgtggcaccttccaatgatgaacaaggcgtactatcaacaatcgaa
 aatgttgttctcggttattccaataaataa

45 Sequence 2472

MIKAIIVDMDGTFLDTNKQFDRNRFETIFKELIDKNIKFIAASGNQFAKLKSI FG DREMF
 FISENGAVIYKGNQLYNYRSFDQYIFQKVNYLNLNQKINNLIICGVKSAYILKETSEAF
 KQDARTYYHQLIEVDSLQTLPPDDYVKIAFNINRQTHPDLDEKLALFKDDIKLVSSGRD
 SIDVIMPNTKGQALSRLKEWQMPASHLMAFGDANNDKMLELAHESYVMANSEDSLSF
 NIASHVAPSNDQGVLSSTIENVVLGYSNK*

50 Sequence 2473

Contig_0763_pos_7023_7691,

is similar to (with p-value 2.0e-20)

>gp:gp|AF012552|AF012552_2 Helicobacter pylori prolipoprotei
 55 n diacylglycerol transferase (lgt) and NADPH-linked flavin n
 itroreductase (rdxA) genes, complete cds. NID: g2564440.
 atgattatgaatcagatgaatcaaacgattattgatgcattccatttttagacatgcgaca
 aaagaatttgaccctacgaaaaaattagtgatgaagattttaatacgaatttttagaaaca
 ggtagattatctccaagttcactaggttttagaaccttggcactttgtagtggttcaaat

aaagaattgagagaaaaattgaaagcctatagttggggagcacaaaagcaacttgataca
gcaagtcactttgtattaatttttgcctgtaagaatgtgacggctcatcacagattacgtg
caacatttacttcgtggcgtcaaaaaatatgaagaaagtacaattccagcagttgaaaat
5 aaatttgatgatttccaagaaagtttccatattgccgataatgaacgaacatttatgac
tgggcgagtaaacacaacatatattgcattagcaaacatgatgacaagtgctgcattacta
ggtatcgactcatgtccaattgaaggatttgatttagataaaagtgactgaaattccttca
gatgaggggtgttttagatacggacaatttggtatttcagttatggtaggctttgggtac
agagacacaagaacctaataacatggcaagttagacaaaacgaagacgacatcattagttgg
attgaataa

10

Sequence 2474

MIMNQMNQTIIDAFHFRHATKEFDPTKKISDEDFNTILETGRLSPSSLGLEPWHFVVQN
KELREKLKAYSWSGAQKQLDTASHFVLIFARKNVTATHDYVQHLLRGVKYEESTIPAVEN
KFDDFQESFHIADNERTLYDWASKQTYIALANMMTSAALLGIDSCPIEGFDLDKVTIELS
15 DEGVLDTEQFGISVMVGFYRAQEPKHGKVRQNEDDIISWIE*

Sequence 2475

Contig_0763_pos_6789_5797,
is similar to (with p-value 0.0e+00)

20 >gp:gp|U31175|SAU31175_1 Staphylococcus aureus D-specific D-
2-hydroxyacid dehydrogenase (ddh) gene, complete cds. NID: g
1644432.

atgacaaaaattatgtttttcggcacaagagcatatgagaaggacatggcattacgttgg
ggaaagaaaaataatatcgatgtcactacatcaacagaacttttaagtgtagatactgtc
25 gatcaattaaaagattatgacgggtgttacaacaatgcagttcggtaattagaacctgaa
gtttaccctaaatttagagtcctatgggtattaaacaaattgcacaacgtacggctggatt
gatagtgatgacttagaacttgcaaaaaacatgaaattattatctcgaatatacctagt
tattcacctgaaacaattgtgtaatttcgggtatctatcgctctgcaactcgtacgaaaa
ttcccaacaattgaaaaacgtgtgcaagcacataatttcacatgggcgtcccctattatg
30 tctcgtccagtaaaaaatatgactgtagcaatcatcggtacagggcgtatttggtgctgca
actggtaaaaatctatgctggttttggtgagagtagttggttatgatgcatacctaata
cattctttatctttcttagaatataaagaacagtagaggatgcaattaaagatgctgat
attatctcattacatgtaccgctaataaagatagttccatttatttgataacaatatg
tttaaaaatgttaaaaaagggtgccgttttagtcaatgccgcaagaggagctgtgataaac
35 acgcctgatttaattgaagcagtaataatggtacattatcaggtgctgccattgacaca
tatgaaaatgaagctaattatttcacatttgattgttcaaatcaaacgattgacgacca
atattattagacctaatagaaatgaaaatatttttagttacacctcatattgcctttttc
tccgatgaagcagtaaaaaatttagtagagggtggtttgaatgcagcattatcagtaatt
aatactggcacatgtgatacgcgattaaactaa

40

Sequence 2476

MTKIMFFGTRAYEKDMALRWGKKNIDVTTSTELLSVDTVQDKDYDGVTMVFQKLEPE
VYPKLESYGIKQIAQRTAGFDMYDLELAKKHEIIISNIPSPETIAEYSVSIALQLVRK
FPTIEKRVQAHNFTWASPMSRPVKNMTVAIIGTGRIGAATGKIYAGFGARVVGYDAYPN
45 HSLSFLEYKETVEDAIKDADIISLHVPANKDSFHLFDNNMFKNVKKGAVLVNAARGAVIN
TPDLIEAVNNGTSLGAAIDTYENEANYFTFDCSNQTIIDPILLDLIRNENILVTPHIAFF
SDEAVQNI.VEGGLNAALSVINTGTCDTRLN*

Sequence 2477

50 Contig_0763_pos_3669_3175,
is similar to (with p-value 3.0e-17)

>pir:pir|JT0409|JT0409 phosphinothricin-N-acetyltransferase
- Streptomyces viridochromogenes >gp:gp|M22827|STMPAT_2 Stre
ptomyces viridochromogenes phosphinothricin N-acetyltransfer
55 ase (pat) gene, complete cds. NID: g295177. >gp:gp|X65195|SV
PTT_3 S.viridochromogenes genes pms, phsA, pat and dea for p
hosphinomethylmalic-acid-synthase, phosphinothricin-tripepti
de-synthetase A, phosphinothricin-N-acetyltransferase and N-
acetylphosphinothricin-tripeptide deacetylase, respectively.

NID: g47997. >gp:gp|A02774|A02774_1 Artificial phosphinothricin resistance gene. NID: g345279. >gp:gp|A02804|A02804_1 S.viridochromogenes phosphinothricin resistance gene. NID: g345154. >gp:gp|A29201|A29201_1 Synthetic DNA for phosphinothricin resistance gene (viral/herbicide resistance) from patent WO9111517. NID: g1248925.

5 atgattagatttgcacgactagaagatcttcaagatattttgacaatttataatgatgcc
atccttaataacaacagctgtttatacgtataagccacaacaattagatgaacgtcttcaa
10 tggatcaatctaaagcaaaaataaacgaacctatatgggtttatgaaaaagaaggga
gtagttgggttttgcacttatggttcctttagacaatggccggcctatttatatactatt
gaacattctatatatgttcatcaacagtacagaggactaggtatcgcttctcaattatta
gagaatttaattcgttacgctaaagaacaaggttatcgccaccattgttgctgggattgat
gcacgaaacatggatagtagtgcgattgcataagaagtttgacttctcacatgcaggtaca
15 attaaaaatgtaggtataaatttgatcgatggctcgatttatcattttatcaatatgat
ttatctgattcataa

Sequence 2478
MIRFARLEDLQDILTIYNDAILNNTTAVYTYKPPQLDERLQWYQSKAKINEPIWVYEKEGK
VVGATYGSFRQWPAYLYTIEHSIYVHQYRGLGIASQLENLIRYAKEQGYRTIVAGID
20 ASNMDSIALHKKFDFSHAGTIKNVGYKFDRWLDLSFYQYDLSDS*

Sequence 2479
Contig_0764_pos_4801_5163,
is similar to (with p-value 2.0e-52)
25 >sp:sp|Q05217|SECY_STACA PREPROTEIN TRANSLOCASE SECY SUBUNIT
. >pir:pir|S30115|S30115 secY protein - Staphylococcus carnosus
>gp:gp|X70086|SCSECY_1 S.carnosus secY gene. NID: g49188

atgattattttatgtagttttaattattgcatttgcataatttttatgcttttgtacaagtt
30 aatcctgaaaaaatggcagataaccttaaaaagcaaggtagtattgtcccaggaattaga
cctggtgaacaaacaaaaaatatattactaaagtactttatagattgacttttgttggt
tcaattttcttagcagctatagctattttacctataattgcgactaaatttatgggctta
ccacaatcaattcaaatgggtggtagcagctcttttgatcgcttattgggtgtagctattgaa
actatgaaaacttttagaagcacaagtcactcaaaaagaatataaaggctttgggtgtaga
35 taa

Sequence 2480
MIYYVLIIAFAYFYAFVQVNPEKMADNLKKQGSYVPGIRPGEQTKKYITKVLYRI.TFVG
SIFLAATAILPIIATKFMGLPQSIQIGGTSLLIVIGVAIETMKTLEAQVTQKEYK.FGGR
40 *

Sequence 2481
Contig_0764_pos_6903_5197,
is similar to (with p-value 2.0e-26)
45 >sp:sp|P54159|YPBR_BACSU HYPOTHETICAL 137.4 KD PROTEIN IN BC
SA-DEGR INTERGENIC REGION. >gp:gp|L77246|BACYACA_6 Bacillus
subtilis (YAC10-9 clone) DNA region between the serA and kdg
loci. NID: g1256615. >gp:gp|Z99115|BSUB0012_144 Bacillus su
btilis complete genome (section 12 of 21): from 2195541 to 2
50 409220. NID: g2634478.

atgaaacgtattaatgaagttggtataccaattatttttggatgaatcagattgataaa
cataatgaagaagaattacatttgaaacttttaaatcaagagtcgaaaaatcaatcaaa
gactgggatatacaacttcaagatacttattacgtttcaagtttgatcatccacagaat
gaaattgacaaaactttcaaatttcttagtattttatggatcaacatcgatgaatcaacagaa
55 gactatgttaatagaaccattcaattcattaccgacgcacaatacatatacattcaaaat
gaaatgcaatctattcttgacacccttcaaatgaagaacaattcgaggaagcatat
attcaatttcaacaaaatcaagaagtcagcgcagaagcacaattgctcaatgactcta
caattatatttaatttttaaacagaagcgtaaagatatattagataatgcttataatg
acgtacgatatcgctgaatctttacggaattatttagaaagcatggcaactgattttaa

gtgaatggattttttaataaaaaggaagaaaaagaagaagaacaaatcaaacgacttaat
 gaggcgaccactcaattgcaagagaaagttaatcaacaagtagcacaaccacttcgtgaa
 gatattgcatTTTTTaaactagattcataaataaacatgctgtgaatgaaaaaataactaaat
 caagaatatgacgtcggtccgctcacttatatcagagctatatcaaactcaaacgagcatt
 5 agcaacacatacgtttttaacattttcagatgaagttataaaaagctttgaataaaaaaata
 gaaaatgagtgcaacaccactatttgaagaagctgtcaatcatgtacaagttaatgaatta
 tcgagtgatgaaaatgaagataggtatgaatatgatagatacattgaacttaacacatta
 aaggattcgcttacatcccacaactacaaacattactatatccatttagacgattcttta
 gataaattaatggaagaacagagactcattttgaacttaaacagaagaaattcaactgct
 10 ta!catcgtaaacatgagacacaacatcgtaacgagtttgttacatctaataagatatt
 aagcgtgcattagatatcgtaaagatgtaccattatttgatcgactaaacaagatattc
 accgataccattctgagactcgataatcaataacaaaagttgggtgttttggtacattt
 agcgaggtgaaaagcagcttaataacgctctactaggtgaaaactatttagtttct
 cctaattccaacaacagcggaacgacggaattatcatatggttaaagagagtcataatcaca
 15 ttaaaatcgaaagaacaattactagaggaagttaatcatgtactagaattttatgaaata
 tcgtttaacacatttagacgactttattgagagtgatttagataagttaaaattgaaacta
 gaaaagaaccaacttgcatattattagtgcaattgagaaacattatgaaatgtacacatct
 atgttagaacattcacttatcacacagtatcgcttgaagaaattaaaaaatggagtgcc
 gaggatgagtgatgctactttcgtagaaaactgtacaccttaagctacctttagattggctc
 20 aagggtaaaatcattattgcccattaa

Sequence 2482

MKRINEVGIPIIIFVINQIDKHNEEITFETFKSRVEKSIKDWDIKLQDTYYVSKFDHPQN
 EIDKLSNFLVFMDQHRESTEDYVNRTIQFITDAQYIYIQNEMQSILDTLQINEEQFEEAY
 25 IQFQQNQEVSAEAQLLNSNQLFNLYLKQKRKDILDNAYIMTYDMRESLRNYLESMTDFK
 VNGFFNKRKKKEEQIKRLNEATTQLQEKVNQQVRQPLREDMSFLTRFINKHAVNEKILN
 QEYDVVPSLISELYQTQTSISNTYVLTFSEVIKALNKKIENESTPLFEEAVNHVQVNEL
 SSDENEDRYEYDRIELNTLKDSLTSNHYKHYYIHLDDSLDKLIGRTETHFELQENSTA
 YHRKHETQHRNEFVTSNQDIKRALDIVKDVPLFDRTKQDITDITLRLDNQITKGVVGTGTF
 30 SAGKSLINALLGENYLVSSPNPTTAATTELSYGKESQITLKSKEQLLEEVNHVLEFYEI
 SFNTLDDFIESDLKLEKNQLAFISAIEKHYEMYTSMLEHSLIHTVSLLEEIKKWSA
 EDEYATFVKTVHLKPLDWLKGKIIIAH*

Sequence 2483

35 Contig_0764_pos_4419_4120,
 is similar to (with p-value 5.0e-52)
 >sp:sp|P43148|SEPA_STAEP EXTRACELLULAR ELASTASE PRECURSOR (E
 C 3.4.24.-) (SEPP1). >pir:pir|A40659|A40659 elastase, SepP1=
 etrxtracellular metalloprotease - Staphylococcus epidermidis
 40 >gp:gp|X69957|SESEPP1A_1 S.epidermis gene for protease. NID
 : g396258.
 atgtctaataccagagcggttttggaacaaccatctcatatgaatgattttgtttatacaaat
 tctgacaacggaggcggtacatacgaattcagggtattccgaacaaagcagcttacaacaca
 attcgtagtatttgtaaacacggttctgaacaaatttattatagagcggttaactgtttat
 45 ttaacttcaaatctgatttccaagatgccaagcatcattacaacaagcagcatttgat
 ttatatggcgacggtattgctcaacaagtaggtcaagcatgggacagtggtggcgtgtag

Sequence 2484

50 MSNPFRFGQPSHMNDFVYTNSDNGGVHTNSGIPNKAAYNTIRSIGKQRSEQIYYRALT
 LTSNSDFQDAKASLQQAFFDLYGDGIAQQVGQAWDSVGV*

Sequence 2485

55 Contig_0764_pos_3606_3145,
 putative peptide of unknown function
 atgaaaaagagtaaacgacaagatttagtaactatgattgttaagcaaaatcacatttat
 aaaaaagcagatattattgattacattgatgatcactttgggtgtacgttatagcatgact
 actattgagagagatttaagagaacttcatatttatcgctgcctgtgaaggcaaatcaa
 tatgaatacaaaattacttacgcaacaatctcaattagactcaagagtaagactaaatgat

tatatagaaacagaaattattaacactatgattaaagaatcgtatataactaataaaagacg
acaccaggtattgacaaaagcataaattattatattgatcagttacaactgaaagagatt
ataggaacgataagtggtaatgatacaattatgattcttacacattcccagtcctatagct
gaatatgtctattacaaaatatttaatacataattattcataa

5

Sequence 2486

MKSKSRQDLVTMIVKQNHYYKKADIIDYIDHFGVRYSMTTIARDLRELHIYRLPVKANQ
YEYKLLTQSQQLDSRVRLNDYIETEIINTMIKESYILIKTTPGFAQSINYIIDQLQLKEI
IGTISGNDTIMILTHSQSIAEYVYYKIFNHNYS*

10

Sequence 2487

Contig_0764_pos_2812_1742,

is similar to (with p-value 0.0e+00)

>gp:gp|Y17554|BLY17554_1 Bacillus licheniformis arcA, arcB,

15

arcC and arcD genes. NID: g3687415.

atgagtatgacaaatggacattcaagttaacagtgaaatcggcaaattaaagacggta
ttacttaaaagaccaggggaaggaattagaaaacttagtgccctgattacttagatggatta
ttgtttgatgatattccatttttaaggtagctcaacaagagcatgattttgtctcaa
gttcttcaagatgaaggaatagaagtgcctttatttagaaaaattagcagcacaaagtata
20 gaagattcaaatgtcagagagcaatttatagatgacgttttagcagaatctagaaaaact
atcctaggtcatgaaaaagaaataaaaaaactcttttcaactttgtcaaataagcatta
attaataagattatggctggcgtacgcaaagaagaatacaacttgaatcgacacatctt
gtagagtacatggatgataaataatccgttttatcttgacccaatgcctaatactatattc
acacgtgatccctcaagcttcaattggtagaggtatgacagtaaatcgtatgttttgaga
25 gcgagacgcagagaatcgattttcatttcataattttaaaacatcatcctagatttaaa
gatgagaatattccctttatgggtggatcgctgactgctcggttcaacatcgaaggtggagac
gaactgggtgttatctaaagatgtacttgcaatagggatatctgaacgtacttctgcacaa
gcaattgaacgttttagcacgacgtatttttaagatccgttatctacttttaaaaggtg
gtggcgattgagattccaactagtcgaacatttatgcacttagatactgtttgtacaatg
30 attgattacgacaaattcactacacattcagcaattcttaaatcagaaggaaacatgaat
atctttattatcgaatatgatgataaagctgaagatatcaaaatccaacattctagtcat
cttaaacaaacattagaagaagtgcctcgatgttgatgaaatcacattaataccaactgga
aatggtgatcatcgacggtgctcgatgaatggatgtccaggtcctggtga

35

Sequence 2488

MSMTNGPIQVNSEIGKLKTVLLKRPKLELENLVPDYLDGLLFDDIPFLKVAQQEHDHFAQ
VLQDEGIEVLYLEKLAAQSIEDSNVREQFIDDVLAESRKILGHEKEIKKLFSTLSNQAL
INKIMAGVRKEEIQLESTHLVEYMDCKYPFYLDPMPLNYFTRDPQASIGRMTVNRMFWR
ARRRESIFISYILKHHPRFKDENIPLWDRDCPFNIEGGDELVLSKDVLAIGISERTSAQ
40 AIERLARRIFKDPPLSTFKKVVAIEIPTSRFTMHLDTVCTMIDYDKFTTHSAILKSEGNNM
IFIEYDPAEDIKIQHSSHLKQTLLEVLDVDEITLIPTGNGDIIDGAREWMSRSI/*

Sequence 2489

Contig_0765_pos_772_1692,

putative peptide of unknown function

45

gtgatttcactactggtgtgtttgttacttttctgttggttcaccttcgccaactttt
tccccgttaaatgggttcttagttgttggtgttgtaattgttttgttcctggttcacct
ttctgtttaacgcgctctttacctggttttaaatcaggattgaattcacgtttcttgcg
aatggaatttcttccgttgacgtgatcgatctccatcaactggaccatattttgtcaca
50 tcatccactgggtgtgactacttcgcctgtatcagggtttttaactcctggtttacct
ggaacgtcctcttggtacctttcggtgcatttggtatcaaatcatccttatggcctggc
ttgatttcttcgccaccatattctgtgatttcatctactggttgtttgttatttttct
gttggttcaccttcgccaaacttttccctgttaaatgggttcttagttgttggtgttgta
attgtttttgttctcgttgcaccttttgtttaacacgctcttcacctggttttaaatca
55 ggattgaattcacgtttcttgcgaatggaatttcttccgttgacgtgatcgatctcca
tcaactggaccatattttgtcacatcatccacaggtggagtaactacttcgcctgtatca
ggatttttaaccccccggttacctggttgcgttggttgactaccttcggtgcatttgg
atcaaattcatccttatggcctggctgatttcttcgccaccataatgaacgatttcatc
cactggttggtttgttatttttctgttggttcaccttcgccaacttttctcctgtatt

aggattgacataagttggtgtgtgtgtgtttcaattcctggttcacctttttggactac
ttttctgtacctgggctaa

Sequence 2490

5 VIISSTGCFVTFVSGSPSPTFSPVNGFLVVGVVIVFVPGSPFCLTRSLPGFKSGLNSRFLS
NGISSVDVIGSPSTGPFVFTSSTGGVTTSPVSGFLTPLPGTSSWLPFGAFGSNSSLWPG
LISSPPYVISSTGCFVIFVSGSPSPTFSPVNGFLVVGVVIVFVPGSPFCLTRSSPGFKS
GLNSRFLSNGISSVDVIGSPSTGPFVFTSSTGGVTTSPVSGFLTPLRLTWLRCLTTFRCIW
10 IKFILMAWLDFFATIMNDFIHWLFCYFFCWFTFANFFSCIRIDISWCCCCFNSWFTFLDY
FECTWG*

Sequence 2491

Contig_0765_pos_3052_2720,
putative peptide of unknown function

15 atgcagaagcacctcaatctgagccaacgaagacagaagaaggaagcaacgcaaaagcag
ctcaatctgagccaacgaaggcagaagaaggaggcaatgcagaagcagctcaatctgagc
caacgaagacagaagaaggaagcaatgcagaagcacctcaatctgagccaacgaaggcag
aagaaggaggcaatgcagaagcacctcaatctgagccaacgaagacagaagaaggaggca
atgcagaagcaccgaatgttccaactatcaaagctaattcagataatgatacacaaacac
20 aatttcagaagcccctacaagaaatgacctag

Sequence 2492

MQKHLNLSQRRQKKEATQKQLNLSQRRQKKEAMQKQLNLSQRRQKKEAMQKHLNLSQRRQ
KKEAMQKHLNLSQRRQKKEAMQKHRMFQLSKLIQIMIHKNHFQKPLQEMT*

25

Sequence 2493

Contig_0765_pos_2621_1443,
putative peptide of unknown function

gtgaatttaaattatagttctccgtttatgtccttattaagcatgcctgctgatagttca
30 tccaataacactaaaaatacaatagatataccgccaactacggttaaaggtagagataat
tacgatttttacggtagagtagatatcgaaagtaatcctacagatttaaattgcgacaaat
ttaacagagataaattatggacagccacctggtacaacaacagctggtgcagttcaattt
aaaaatcaagtttagttttgataaagatttcgactttaacattagagtagcaaaacaatcgt
caaagtaatacaactggtgcagatggttggggctttatgttcagcaagaaagatggggat
35 gatttcctaaaaaacggtggtatcttacgtgaaaaaggtagacacctagtgagctggtttc
agaattgatacaggatattataataacgatccattagataaaatacagaaacaagctggt
caaggctatagagggtatgggacattttgttaaaaatgactcccaaggtaataacttctaaa
gtaggatcaggtactccatcaacagattttcttaactacgcagataataactactaatgat
ttagatggtaaatccatgggtcaaaaattaaataatgttaatttgaaatataatgcttca
40 aatcaaaacttttacagctacttatgctggtaaaaacttgacggctacgttatctgaatta
ggattgagtgcaactgatagttacaatttttttagttacatcaagtcaatatggaaatggt
aatagtggtacatacgaagtggtggttatgagagctgatttagatggtgcaacattgaca
tacctcctaaagcagtcgatggagatccaattatatcaactaaggaaataccatttaaat
aagaaacgtgaatttgatccaaacttagccccaggtagcagaaaaagtagtccaaaaagg
45 gaaccaggaattgaaacaacaacaacaccaaacttatgtcaatcctaatacaggagaaaaa
gttggcgaagggtgaaccaacagaaaaataaacaacaacacagtggtgaaatcgttcat
tatggtggcgaagaaatcaagccaggccataaggatgaatttgatccaaatgcaccgaaa
ggtagtcacaacaacgcaaccaggtaagccggggggttaa

50 Sequence 2494

VNLNYSSPFMSLLSMPADSSSNNTKNTIDIPPTTVKGRDNYDFYGRVDIESNPTDLNATN
LTRYNYGQPPGTTTAGAVQFKNQVSFDKDFDNIRVANNRQSNTTGADGWGFMESEKDG
DFLKNGCILREKGTPSAAGFRIDTGYNNNDPLDKIQKQAGQGYRGYGTfVKNDSQJNTSK
VSGTGPSTDFLNYADNTTNDLDGKFHGQKLNNVNLKYNASNQFTTATYAGKTWTATLSEL
55 GLSPTDSYNFLVTSSQYNGNSGTYASGVMRADLDGATLTYPKAVDGDPIISTKEIPFN
KKREFDPNLAPGTEKVQKGEPIETTTTPTYVNPNTGEKVGEGETEKITKQPVDEIVH
YGEEIKPGHKDEFDPNAPKGSQTTQPGKPGG*

Sequence 2495

Contig_0765_pos_0_1408,

putative peptide of unknown function

gtggatgatgtgacaaaatattggtccagttgatggagatccgatcacgtcaacggaagaa
 attccattcgacaagaaacgtgaattcaatcctgatttaaaaccaggtgaagagcgtgtt
 5 aaacaaaaaggtgaaccaggaacaaaaacaattacaacaccaacaactaagaaccatta
 acaggggaaaaagttggcgaaggtgaaccaacagaaaaataacaaaacaaccagtagat
 gaaatcacagaatatgttgccgaagaaatcaagccaggccataaggatgaatttgatcca
 aatgcaccgaaaggttagccaagaggacgttccaggtaaaccaggagttaaaaaccctgat
 acaggcgaagtagtcacaccaccagtggtgatgatgtgacaaaatattggtccagttgatgga
 10 gatccgatcacgtcaacggaagaaattccattcgacaagaaacgtgaattcaatcctgat
 ttaaaaccagggtaaagagcgcgttaaacagaaaaggtgaaccaggaacaaaaacaattaca
 acaccaacaactaagaaccattaacaggggaaaaagttggcgaaggtgaaccaacagaa
 aaagtaacaaaacaaccagtagatgaaatcacagaatatgttgccgaagaaatcaagcca
 ggccataaggatgaatttgatccaaatgcaccgaaaggttagccaagaggacgttccaggt
 15 aaaccaggagttaaaaatcctgatacaggcgaagtagttactccaccagtggtgatgtg
 acaaaaatattggtccagttgatggagatccgattacgtcaacggaagaaattccgtttgat
 aaaaaacgcgaatttgatccaaacttagcgcagggtacagagaaagtcgttcaaaaaggt
 gaaccaggaacaaaaacaattacaacaccaacaactaagaaccattaacaggggaaaaa
 gttggcgaaggtgaaccaacagaaaaagtaacaaaacaaccagtggtgaaatcgttcat
 20 tatgttgccgaagaaatcaagccaggccataaggatgaatttgatccaaatgcaccgaaa
 ggtagccaagaggacgttccaggtaaaccaggagttaaaaaccctgatacaggcgaagta
 gttactccaccagtggtgatgatgtgacaaaatattggtccagttgatggagatccgattacg
 tcaacggaagaaattccgtttgataaaaaacgcgaatttgatccaaacttagcgcagggt
 acagcgaagagtcgttcaaaaaggtgaaccaggaacaaaaacaattacaacaccaacaact
 25 aagaaccattaacaggggaaaaagttggcgaaggtgaaccaacagaaaaagtaacaaaa
 caaccagtggtgaaatcgttcaTATCG

Sequence 2496

VDDVTKYGPVDGDPITSTEEIPFDKKREFNPDLPGEERVKQKGEPGKTKITPTTKNPL
 30 TGEKVGEGETEIKTKQPVDEITEYGGEEIKPGHKDEFDPNAPKGSQEDVPGKPGVKNPD
 TGEVVTTPVDDVTKYGPVDGDPITSTEEIPFDKKREFNPDLPKGERVKQKGEPGKTKIT
 TPTTKNPLTGEKVGEGETEIKTKQPVDEITEYGGEEIKPGHKDEFDPNAPKGSQEDVPG
 KPGVKNPDTGEVVTTPVDDVTKYGPVDGDPITSTEEIPFDKKREFDPNLAPGTEKVVQKG
 35 EPGTKTITPTTKNPLTGEKVGEGETEIKTKQPVDEIVHYGGEEIKPGHKDEFDPNAPK
 GSQEDVPGKPGVKNPDTGEVVTTPVDDVTKYGPVDGDPITSTEEIPFDKKREFDPNLAPG
 TEKVVQKGEPGKTKITPTTKNPLTGEKVGEGETEIKTKQPVDEIVHIX

Sequence 2497

Contig_0766_pos_1949_3403,
 40 is similar to (with p-value 1.0e-62)
 >sp:P14567|DP3A SALTY DNA POLYMERASE III, ALPHA CHAIN (EC
 2.7.7.7). >pir:pir|A45915|A45915 DNA-directed DNA polymeras
 e (EC 2.7.7.7) III alpha chain - Salmonella typhimurium >gp:
 gp|M29701|STYDNAE_1 S.typhimurium polymerase III polymerase
 45 subunit gene, complete cds. NID: g153951.
 atggaagaaataaccaacttatataaccctagacataatcctaaccaatttgcttattta
 catccagatttagaaccaatcttaaaaaacacatatggtgttatcatttatcaagaacaa
 ataatgctaataagcaagtaagttgctggttttagttatggtgaagcagatattttaaga
 agggcaatgagtaaaaaagaatcgtgcaatcttagaaagtgagcgtcaacatttcattgat
 50 ggtgcaaaaaataacggttacgatgaacagataagtaagcaaatttttgatttaatactt
 aagtttgcagattatgggttccacagtgccatgctgttagttactcaaaaattgcatac
 attatgagctattttaaaagtgcactatcctcattatttttatgcaaatatcttgagtaat
 gtaataggaagtgaaaaaagactgcagctatgattgacgaagctaagcaccaaagaatt
 agcatcttgctctcccaatatttaaaagtcattggtattataaggcaagtaataaaagga
 55 atatatctgtcttttaggtacaattaaaggaattggatatcaaagcgttaaatattatt
 gatgaacgtcagcagaatggaccttatagagatttctttgatttttcaagacgtatacca
 aaaagggtagaaaaatagaaaattacttgagtctcttatcttagtaggcgattcgacact
 ttggcaaaactagagcgacattattacaagcaattgatcaagtattagatttgaattct
 gatgttgagcaagatgaaatgcttttcgatcttttaactcctaacaatcgtatgaagaa

aaagaggaactacctgatcaattattaagtgattatgaaaaagaatacctaggattctat
 attagtaaaccatccagttgaaaagaaatttgaaaagaacaataatttaggcataatttcaa
 ttgtctaattggaagtcactaccaacctaacttgttcaatttgaccatatcaaacaaata
 5 agaacgaagaatgggtcaaaatatggcatttgaacgatgaatgatggaagaacgatgatg
 gatggagtgattttccagataagtttaaaaaatacgaacttctatttcaaaggacag
 atgtatatcgatttaggttaaatttgaaaagcgtaaccaacaatacgcaacttatcatcaat
 caactttttgaagttgaagcgatgatgcaaacaaaattgtctaatcgaaaaaagttatt
 ttacgtaatgtaacacatctagaaccacaatttgaacattcaaaagtagaatctaataa
 caacatgcattaaatatttatggttttgacgaaagtgcaataaagatgacaatgttggga
 10 caaattgaacgtcaacgtcaaaattttgatctattaatacaaaacttattcgccagctgat
 attagattttatttaa

Sequence 2498

MEEIPTYITRRHNPQFAYLHPDLEPILKNTYGVIIYQEIQIMLIASQVAGFSYGEADILR
 15 RAMSKKNRAILESERQHFDGAKNNGYDEQISKQIFDLILKFADYGFPRAHAVSYFKIAY
 IMSYIKVHYPHYFYANILSNVIGSEKKTAAMIDEAKHQIRISILPPNINQSHWYK-SNKG
 IYLSLGTIKGIGYQSVKLIIDERQONGPYRDFDFSRRIKRVKNRKLLESLLVGAFTD
 FGKTRATLLQAIQVLDLNSDVEQDEMLFDLLTPKQSYEEKEELPDQLSDYEKEYLGFY
 ISKHPVEKKFEEKQYLIGFQLSNGSHYQPIVQFDHIKQIRTKNGQNMFAVTMNDGRTMM
 20 DGVIFFDKFKKYETSISEKQMYIVLGKFEKRNQMQQLIINQLFEVEAYEQTKLSNSKKVI
 LRNVTHLEPQFEHKSVESENEQHALNIYGFDESANKMTMLGQIERQRQNFLLIQTYS PAD
 IRFI*

Sequence 2499

25 Contig_0766_pos_3567_4364,
 is similar to (with p-value 7.0e-79)
 >sp:sp|P54537|YQIZ_BACSU PROBABLE AMINO-ACID ABC TRANSPORTER
 ATP-BINDING PROTEIN IN BMRU-ANSR INTERGENIC REGION. >gp:gp|
 D84432|BACJH642_255 Bacillus subtilis DNA, 283 Kb region con
 30 taining skin element. NID: q2627063. >gp:gp|Z99116|BSUB0013_
 107 Bacillus subtilis complete genome (section 13 of 21): fr
 om 2395261 to 2613730. NID: q2634723.
 atgattgtactattgttctacataatttactagtaaatcggtgtatattttaaagtattcag
 ttttcttttcagccagtgattaatattaaaaatttaataaaaaatttgagcaaatgaa
 35 gtaattgagagatatattaacttactgttgaaaagggatgaagtggttgcaataattggacca
 tctggaagtggtgaaaagcactttactccgttgatgaatttggtagatgtaccttcaaaa
 ggtaaagttatatttgaagataatgaattaaactcaacataatgttcatttagataattta
 cgacaaaaaatgggtatgggtatttcaaaattttaatttatttcctcataaaaaaggtcatt
 gaaaatgtaattgttgaccactttttattacataaagatagtaaagatcaattaaaagag
 40 aaagcttttatatttacttgaaaagtggggcttaaaagacaaagcagattcatatccta
 caactgtctggaggtcaaaaacaaagagttgctattgcaagagctttggcaatggaacct
 gatgttatgttatttggatgaaccaacatctgcacttgatcctgaagtagtaggggatgtt
 ttaaaagtaattgagacaattagcgaatgaaggatgacaatggtgattgtcacgcatgag
 atgaactttgctaaagaaataagtgataaagtagtatttatggccgatggtgttgggtt
 45 gaatctggtacaccacaaaacatatttgaaaatcctcagcacagtcgaactgaaaatttt
 ttatcacgagtggttataa

Sequence 2500

MIVLLFYIFTSKSCIFKMIQFSFQPVINIKNLNKKFGANEVLRDINLTVEKGEVVAIIGP
 50 SGSGKSTLLRCMNLLOVPSKGVIFEDNELTQHNVLHDLNRQKMGVMVFQNFNLFPHKKVI
 ENVMLAPLLLHKDSKQDLKEKALYLLEKVGLKDKADSYPNQLSGGQKQRVAIARALAMEP
 DVMLFDEPTSLDPEVVGDLKVMRQLANEGMTMVIIVTHEMNFKEISDKVVFMDGVVV
 ESGTPQNI FENPQHSRTENFLSRVL*

Sequence 2501

55 Contig_0766_pos_4706_5149,
 is similar to (with p-value 1.0e-20)
 >pir:pir|S05373|S05373 ctc protein - Bacillus subtilis (frag
 ment)

atggtttcagattatcaatttgatccattaaaaaaccaaatcactcatattgacttttta
 gcaatcaacatgagtgaagaacgtactgttgaagtacaagttcaattagttggtgaagct
 gtaggtgctaagaaggcggcgtagttgaacaaccattattcaacttagaagttacagct
 acacctgaaaatattcctgaaactatcgaagtagatatcagtgaattacaagttaatgac
 5 agcttagcagtttctgatattaaaatctctggtgatttcacaatcgaaaataatccagaa
 gattctatcgtaacagtagttcctccaacagatgaaccttctgaagaagaagttgaagct
 atggaaggcgaatcagcaactgaagaaccagaagttgttggtgaagacaaagaagacgat
 gaagaagaaaataaagaagactaa

10 Sequence 2502
 MVSDYQFDPLKNQITHIDFLAINMSEERTVEVQVQLVGEAVGAKEGGVVEQPLFNLEVTA
 TPENIPETIEVDISELQVNDLAVSDIKISGDFTIENNPEDSIVTVVPPTDEPSEEEVEA
 MEGESATEEPEVVGEDKEDDEEENKED*

15 Sequence 2503
 Contig_0766_pos_5387_0,
 is similar to (with p-value 4.0e-49)
 >sp:sp|P37470|SP5C_BACSU PROBABLE PEPTIDYL-TRNA HYDROLASE (E
 C 3.1.1.29) (PTH) (STAGE V SPORULATION PROTEIN C). >gp:gp|D2
 20 6185|BAC180K_116 B. subtilis DNA, 180 kilobase region of rep
 lication origin. NID: g467326. >gp:gp|Z99104|BSUB0001_53 Bac
 illus subtilis complete genome (section 1 of 21): from 1 to
 213080. NID: g2632267.
 gtggaggtaacaataatgaaatgcattgtcggctcttggaacatttggtaaacgttttgaa
 25 ttaacaagacataatattggtttcgaagttgtcgatgatattctagaacgccaccaattt
 actttagacaaacaaaaatttaaagggtgcataactattgaacgtttaaacggcgaaaaa
 gtatrat.tatttgagccaatgaccatgatgaacttatctggtcaagctgtagccc:ttta
 atggattattataatgtcgatgttgaagatttgatcgttttatatgacgatttagattta
 gaacaaggacaagtgcgtctgcgcaaaaggaggagtcaggcggtcataatggtatgaaa
 30 tcgataattaaaatgcttggtacagatcaatttaaactgattcgaaattggtgttgccgt
 ccaacaaatgggatgtctgttcggactatgttttacaaaaattttcaaaaagaagaatg
 atcattatggaaaaggtaattgaacattctgcaagagctgtagaatcttttattgaaagt
 tctcgttttgatcatgttatgaatgaa

35 Sequence 2504
 VEVIMKCIIVGLGNIGKRFELTRHNIGFEVDDILERHQFTLDKQKFKGAYTIERLNGEK
 VLFIEPMTMMNLSGQAVAPLMDYYNVVDVEDLIVLYDDLLEQQQVRLRQKGSAGGHNGMK
 SIIKMLGTDQFKRIRIGVGRPTNGMSVPDYVLQKFSKEEMIIMEKVIEHSARAVESFIES
 SRFDHVMNE

40 Sequence 2505
 Contig_0766_pos_1682_1317,
 putative peptide of unknown function
 atgtatagtgaaaaagaaatcatcacgaaaagtcgaaagtttagcagagaaaattggaaaa
 45 ctagaagttgttcaagattatcataatgtagaaaaacaaattcataataatcaagcaata
 aaacaaaagatgaatcgtttgaaagcgcaacaaaaacaatcggttaattttcaaaattat
 ggaaaacaaaatgcactcgagcaatctgaagttaaaattcagaatctaaaagatgaaatt
 aatgaattacattgttgaagaatttcggttcagcacaatatgaagcgaatgatttacta
 caaatgatggtcaaaaacaatggaagatagactcaatgaatataataaaaaagaacacaat
 50 gaataa

Sequence 2506
 MYSEKEIIRKVESLAEKIGKLEVVQDYHNVEKQIHNNQAIKQKMNRLKAQQKQSVNFQNY
 GKQNALEQSEVKIQNLKDEINELPIVEEFRSAQYEANDLLQMMVKTMEDRLNEYNKKEHN
 55 E*

Sequence 2507
 Contig_0766_pos_1293_799,
 putative peptide of unknown function

atgaatattagaaagttgaccatcacagcattcttaatagctattaatgtcgtgtaagt
 agtttaattgtcattcccttaggtccaattaaagccgacccgttcaacattttgtaa
 gtattatgtgctgtatttgttggaccatggtatggttttagcgcaagcttttattcttct
 gtactacgaatttcatttggaaactggaagcgcatgttctttccaggaagtatgataggt
 5 gtcttactttccagtcgttttatatgtataggaagcatattttatggcttcagttggt
 gaagtattagggaactggtgttattggtagtttaatgtgtataaccttagcatggtttta
 ggacttcaagatttctttattaaaccattaatgcttatgttcatagtatcaagttttatt
 ggggctttaattagttatatattgcttatttttgaagaagaggtttactagataga
 tttacaaaaattaa

10

Sequence 2508

MNIRKLTITAFLIAINVVLSSLIVPLGPIKAAPVQHFNVLCAVFGVPWYGLAQAFISS
 VLRISFGTGSFAFAFGSMIGVLLSSLFYMYRKHFIMASVGEVLGTGVIGSLMCIPLAWFL
 GLQDFFIKPLMLMFIVSSFIGALISYILLIILKRRGLLDRFNKN*

15

Sequence 2509

Contig_0766_pos_0_613,

is similar to (with p-value 4.0e-50)

>sp:sp|P49849|MUTS_BACSU DNA MISMATCH REPAIR PROTEIN MUTS. >
 20 gp:gp|U27343|BSU27343_2 Bacillus subtilis spore coat protein
 (cotE) gene, partial cds, and mismatch repair recognition p
 roteins (mutS) and (mutL) genes, complete cds. NID: gl002518

25

gtgaaaatagaaatggctaacattacaccaatgatgcaacaatatttaaagataaaatct
 gaatatgatgattgtttgctatttttttagactcggagatttctatgaaatgttctttgat
 gatgctaagaagcatcaagagtacttgaaataacattgacgaaaagagatgctaaaaaa
 gaaaatcctattccgatgtgtggcggtaccatatcattctgctgataattacattgaaaca
 ttgattaataatgggtataaggtcgctatatgtgaacaaatggaagatccaaagcaaaca
 aaaggaatgggttagaagagaagttgtaagaatcatcacaccaggaactgttatggatcaa
 30 aatggtatggatgaaaagaaaaataattatattttaagttttatcgaaaatgaagaattt
 ggattatgctattgtgatgtttctacaggcgaactcaaagtaactcatttcaaagataca
 gcaaccttgcttaatgagattacaacaattaatccaatgaaatcgtcataaagcaagct
 ctatctgaagaattaaaaagacaaatcaacatgataactgagacgattactgttcgcgaa
 gatataTATGTTT

35

Sequence 2510

VKIEMANITPMQYLLKIKSEYDDCLLFFRLGDFYEMFFDDAKEASRVLEITLTKRDAKK
 ENPIFMCCVPYHSADNYIETLINNGYKVAICEQMEDPKQTKGMVRREVVRITPGTVMDQ
 NGMDEKKNYILSFIENEEFGLCYCDVSTGELKVTHFKDTATLLNEITTINPNEIVIKQA
 40 LSEELKRQINMITETITVREDIYVX

Sequence 2511

Contig_0767_pos_5389_5703,

is similar to (with p-value 1.0e-24)

>gp:gp|AF006075|AF006075_3 Bacillus subtilis acetoin dehydro
 45 genase enzyme system gene cluster, ribosomal protein L6-like
 protein gene, partial cds, TPP-dependent acetoin dehydrogen
 ase, E1 alpha-subunit (acoA), TPP-dependent acetoin dehydrog
 enase, E1 beta-subunit (acoB), dihydrolipoamide acetyltransf
 50 erase (acoC) and dihydrolipoamide dehydrogenase (acoL) genes
 , complete cds, and regulatory protein (acoR) gene, partial
 cds. NID: g2957145. >gp:gp|Z99108|BSUB0005_76 Bacillus subti
 lis complete genome (section 5 of 21): from 802821 to 101125
 0. NID: g2633055. >gp:gp|D78509|D78509_4 Bacillus subtilis Y
 55 fjG-YfjR genes, complete cds. NID: g2780390.

atgagttcagcaattggacgtagccctgtcgcagcagctcctactgccgtgctcaatgta
 atgtgttcagcaattggtgtatcgattacacgtttacgactatattttttgcaagtcct
 ttgtgtacaccaaatacaccacgaatgtatcgatcattttgatgtggtctactttcgca
 ccacctgagacatcagtagcaattaaaatgacatcctcatcttttccatagattggtca

atagcttcattaatagccccataaatgttaacttgcgttcttcactcatcgtaaaaacc
tccttatgtggttag

Sequence 2512

5 MSSAIGRSPVAAAPTAVLNVLMCSAIGVSITRLRLYFFASPFVTPNTPPNVSSSLMWSTFA
PPETSVPIKMTSSSFSDWSIASLIAPINVLNRSSLIVKTSLCG*

Sequence 2513

Contig_0767_pos_5685_4645,
10 is similar to (with p-value 0.0e+00)
>gp:gp|AF006075|AF006075_3 Bacillus subtilis acetoin dehydro-
genase enzyme system gene cluster, ribosomal protein L6-like
protein gene, partial cds, TPP-dependent acetoin dehydrogen-
ase, E1 alpha-subunit (acoA), TPP-dependent acetoin dehydrog-
15 enase, E1 beta-subunit (acoB), dihydrolipoamide acetyltransf-
erase (acoC) and dihydrolipoamide dehydrogenase (acoL) genes
, complete cds, and regulatory protein (acoR) gene, partial
cds. NID: g2957145. >gp:gp|Z99108|BSUB0005_76 Bacillus subti-
lis complete genome (section 5 of 21): from 802821 to 101125
20 0. NID: g2633055. >gp:gp|D78509|D78509_4 Bacillus subtilis Y
fjG-YfjR genes, complete cds. NID: g2780390.
atgagtgagaacgcaaggttaacatttatgggggctattaatgaagctattgaccaatct
atggaaaaagatgaggatgtcatttttaattggtactgatgtctcaggtggtgcaaaagta
gaccacatcaaagatgacgatacattcgggtggtgatttgggtgaacaaaaggacttgca
25 aaaaaatatagtcgtaaacgtgtaatcgatacaccaattgctgaacacattacattgagc
acggcagtaggagctgctgcgacagggctacgtccaattgctgaactcatgttcaacgac
tttattgatttgggttagatccaatttttaaatcaaggggcaaaaatgagatatatgttt
ggtggaaaagccaaaatcccactagttgtacgtactcttcatggagcaggggcaagcgct
gctgcacagcactctcagtcctttatataatgtttgcagcaattccaggagttaaagt
30 gttgttccatctaattccatgatgacgaagggtctactgatgtcagctattcaaggagac
aatcttgttgtcttttcagaagataaaacattattaggacaaaaaggtaatgttcctgaa
gaaccttatactatagaaattggtaaaagccaatgtgacgcgtgaaggtgacgatttaaca
attgtggctatttgaaaaatggttagctgtagcgaagaaactgctgaaaaacttgacagaa
gaccaagtatcagttgaggtcatcgatttacgctcaggtgcaccatgggatcaagaaaca
35 gtttttagattctgtgaagaaaacgggtcgcttaattgttattgacgaatctaataccacag
tgtaacattgctggagacgttgcttcagtgattggagatgtaggatttgattacttagat
ggtccaattaagaaagtaccgcaccagacactcctgtaccatttgacgcaacttagag
gcggcatatatgccgaatgctgataaggtatttagacattgcatctgaattaattgatgat
ttaaaaaaggctaacgcataag

Sequence 2514

MSEERKLTFMGAINAIDQSMKEDDEVILIGTDVSGGAKVDHIKDDDTFGGVFGVTKGLA
KKYSRKRVIDTPIAEHITLSTAVGAAATGLRPIAELMFNDFIGFLDPILNQGAKMRYMF
GGKAKIPLVVRTLHGAGASAAAQHSQSLSLYNMFAAIPGVKVVVPSNPYDAKLLMSAIQED
45 NLVVFSEDKTLGQKGNVPPEPYTIEIGKANVTREGDDLTIVAIGKMVAVAEETAEKLA
DQVSVEVIDLRVSPWDQETVLDQSVKKTGRLVIDESNPQCNIAGDVASVIGDVGFDYLD
GPIKKVTAPDTPVPFAANLEAAYMPNADKVLDIASELIDDLKKANA*

Sequence 2515

Contig_0767_pos_4613_3354,
50 is similar to (with p-value 3.0e-72)
>gp:gp|Z99108|BSUB0005_77 Bacillus subtilis complete genome
(section 5 of 21): from 802821 to 1011250. NID: g2633055. >g
p:gp|D78509|D78509_3 Bacillus subtilis YfjG-YfjR genes, comp-
55 lete cds. NID: g2780390.
atgccaaagcttggaatgacaatgaaagaggggaactgttgaagagtgggtttaaatcagag
ggtgacaccgtaaaacaaggagagagattgttacaataagctctgaaaaattaaccaac
gatgttgaagcgccggcagtggtgacattgttagaaattaaagtgaagccggagagaagat
gcagaggttaaaagcgggtattaggtataattggagaagaaggggaagctattgataaagat

gaagatgatttagcatcagaaaaagtaaaagaagacaacgagcatgagaaggaaa:gcaa
 gaagttaaagatacatcacacagctcttccgataataaagataattcgctaaaaagcgca
 gcacgagaaaagaatctttatctcaccctcgacgtaatatggctgaggataaaggatta
 gacattaacaagataaaaaggcacaggcggtaatcatcgtattacaaaactagatattcaa
 5 cgtgttgaagcaaatgggtacgactatgctagtgtacgacatctaataagatacaagt
 catgttccaacacagactgtggatacaagtgcgattgggtgaaggattgaatcctatgcgt
 caacgtattgctcaaacatgagacaaagtcttaatagtactgctcaattaacattacat
 cgtaaggttgtatgcggatcgcttgctagatttcaaagacagattagctacggaacttaaa
 caagcagatcaagatgttaaatgaactgttactacattattagctaaagcagtagtgctt
 10 gcacttaagaatatggggcaatgaatgctcgctatgaacaaggcgagttactgagtat
 gaagatgttcatcttaggaatcgcaacgtctctagatgaaggccttatgggtgccagtgatt
 aatcatgcagatacaaaaagtatcggcacttttagcccatgaaattaaatcatcggtgag
 gctgttcgggaaggaaacacaggagcagtaacaattagaggagcaacatttacaattact
 aatatgggtgctagtgttatagaatactttacaccaattttaaatttaggtgaaacaggt
 15 attctaggcgcttggtgctttaactaaagaagtcgtgctagaagcgagataacattaaaca
 gtttcaaaaattcctttaagcttgacatttgatcatcaaatttttagatgggtgcaggtgcg
 gccgattttcttaaagtactagctaaatatatcgaaaacccttatttattaatgttatag

20 Sequence 2516
 MPKLGMTMKEGTVEEWFKSEGDTVKQGESIVTISSEKLTNDVEAPASGTLLEIKVQAGED
 AEVKAVLGIIGEEGEAIDKDEDDLASEKVKEDNEHEKETQEVKQTSQQSSDNKDNSPKSA
 ARERIFISPLARNMAEDKGLDINKIKGTGGNHRITKLDIQRVEANGYDYASDTTSNEDTS
 HVPTQTVDTSIAGEGLNPMRQRI AQNMQRSLNSTAQLTLHRKVDADRLLDKDRATELK
 25 QADQDVKLTVTTLLAKAVVLALKEYGAMNARYEQGELTEYEDVHLGIATSLDEGLMVPVI
 NHADTKSIGTLAHEIKSSAEAVREGNTGAVQLEGATFTITNMGASGIEYFTPIILNLGETG
 ILGVGALTKEVVLEADNIKQVSKIPLSLTFDHLQILDGAGAADFLKVLAKYIENPYLLML*

30 Sequence 2517
 Contig_0767_pos_2671_2159,
 putative peptide of unknown function
 atgaaaagggttgcaaaagcatttgctcgtaagtgggtattacttttaggtgcagttttagggt
 ttaaacgtaacagagcataatgggtgtatctaataagcaaaaggcacaacagcacacaggt
 35 tactgggtataaatataatgggtatactgcacgggtggcgactttgtacttagcaattca
 ttttatcaagggtttaaagctggaaacggttacatttaattgggtattaaggtaaatcaaaaa
 tatgaatctaagactgctactaaaaaaatatacgatcagacatttcaacaaattaatgga
 aataaagcaaacacgtacaaatttaaaattgcttccagaactgttacttttagatcaagtt
 aaacaaaagtatggaaaaaattataattatcagccgtcattatctaaaaacaaaacaagt
 40 aagacagatggcttgtagcgttatcaagtcggaaaaggaaacatcgttttccacgtt:aaa
 gatgggtatgtcacaaagtgcacattgtcataa

Sequence 2518
 MKRFAKAFVVSGITLGAVLGLNVTEHNGVSNEAKAQTASYSYKYNGYTASGGDFVLSNS
 45 FYQGLKAGNVTFNGIKVNQKYESKTATKKIYDQTFQQINGNKANNVQFKIASRTVTLQV
 KQKYGKNYQPSLSKNKTSKTDGLYGYQVGKGNIVFHVKDGIVTSATLS*

Sequence 2519
 Contig_0767_pos_1876_1568,
 putative peptide of unknown function
 atgagtgagaatatcattaataataaaatgatagagaatggtagacctgtaataactgat
 gcagtttgtaaaccttttaagaccacctgcaattgtcatagcaactgaaatcgccccaata
 agtagaccctcataggactttatgttttaaagttggattgatagaaccacctgtagccata
 cctgcaacaatatgcgtagtgtgaatcagcacttgttacaataaagatgaaaataagtaca
 55 atagctaaggaaactagtgtacttccgacaatggaaaatgtgataataattcaaacaaaggca
 acagtataa

Sequence 2520
 MSENIINNMIENGRPVITDAVCKLLRPPAIVIAIEIAPISRPHRTLCKFKVGLIEPPVAI

PATICVVESALVTIKMKISTIAKELVTS DNGKCDNNSNKATV*

Sequence 2521

Contig_0768_pos_2298_2642,

5 putative peptide of unknown function
gtgacaaaccggaggaaggtggggatgacgtcaaatacatcatgccccttatgatttgggc
tacacacgtgctacaatggacaatacaaaaggtagcgaaccgaggtcaagcaaattcc
cataaagttgttctcagttcggattgtagtctgcaactcgactatatgaagctggaatcg
ctagtaatcgtatgacagcatgctacggtgaatacgttcccgggtcttgtacacaccgcc
10 cgtcacaccacgagagtttgaacacccgaagccggtggagtaaccatttggagctagcc
gtcgaaggtgggacaaatgattgggggtgaagtcgtaacaaggtag

Sequence 2522

VTNRRKVGMTSNHHAPYDLGYTRATMDNTKGSETARSSKSHKVVLSSDCSLQLDYMKLES

15 LVIVDQHATVNTFPGLVHTARHTTRVCNTRSRWSNHLELAVEGGTNDWGEVVTR*

Sequence 2523

Contig_0768_pos_7279_6914,

is similar to (with p-value 1.0e-27)

20 >gp:gp|U96108|SCU96108_3 Staphylococcus carnosus (3R)-hydrox
ymyristoyl acyl carrier protein dehydrase homolog (fabZ) gen
e, partial cds, YwpF homolog, single-strand binding protein
homolog (ssb), SceD precursor (sceD), SceA precursor (sceA)
and SceE precursor (sceE) genes, complete cds, and TenA homo
25 log (tenA) gene, partial cds. NID: g2735509.
atgacgaaagatgctcaaatctatgaaaaggaggataacaaaatagcaacattttgcgtt
gctacagaaagaaactacaaagatgatattaatgaaatttctacagattatttactttgt
aaagctttcggtaaaactgccactaatattgaaaagtacactagtcaaggtacttttagta
ggataactggacaaatgcgttcaagaaaatatgaaaaagaaggccaaacacattttgtt
30 acagaattatacgttgaaacaataaaattcatgtccccaaaaataaaaaaatgaaact
ccctctgataatcaatttgaaaacaacacttatcaacctgatgatttagaaataattcat
attttaa

Sequence 2524

35 MTKDAQIYEKEDNKIATFCVATERNYKDDINEISTDYLLCKAFGKTATNIEKYTSQGT
LVTGQMRSRKYKEGQTHFVTELYVETIKFMSPKNNKNETPSDNQFENNTYQPDLEIIH
I*

Sequence 2525

40 Contig_0768_pos_6526_5867,

is similar to (with p-value 5.0e-49)

>gp:gp|U96108|SCU96108_4 Staphylococcus carnosus (3R)-hydrox
ymyristoyl acyl carrier protein dehydrase homolog (fabZ) gen
e, partial cds, YwpF homolog, single-strand binding protein
45 homolog (ssb), SceD precursor (sceD), SceA precursor (sceA)
and SceE precursor (sceE) genes, complete cds, and TenA homo
log (tenA) gene, partial cds. NID: g2735509.
atgaaaaaacatttagttgcatcatcattagctataggactaggcgtttagcaggtaac
gcaggtcatgacgcacatgcaagtgagactacaaatgttgataaagcagagttagctcaa
50 aaagcgttaactaatgatcaatcactaaatgaaagccctgttcaagaagggtgcttataat
attaattttgattacaatggtaattcatatcattttgaatctgatggttctacttggagt
tgyaatttgaatcaacaacaatgctacgcaacctgttcaaccaagccaatctcagta
gctacacaacaacaacctgtacaagtatcagcaccacaaaatgagcaaacagcacacaacca
caacaacaatcaacatctacaagtcaaaacttcttcaagtaaaagcttcaagtggttcac
55 gtaaatgtaattcacatttacaacaattgctcaacgtgaatcaggtggcgatattcat
gcaataaaccatcttcaggtgcagcaggtaaatatcaattcttacaatcaacttgggat
tctgttgctccaagtcaatataaagggtttcacctgcaaaagctccagaaagcgtacaa
gaccgagcagcagtaaaattatataaacttggtggcccaggtcattgggtaactgcataa

Sequence 2526

MKKTILVASSLAIGLVVAGNAGHDAHASETTNVDKAEALQKALTNDQSLNESPVQEGAYN
 INFOYNGNSYHFESDGSTWSWSYESTNNATQPVPQPSQSQVATQQQPQVVSAPQNECTAQP
 5 QTKSTSTSQTSSSKASSGSSVNVNSHLQQIAQRESGGDIHAINPSSGAAGKYQFLQSTWD
 SVAPSQYKGVSPAKAPESVQDRAAVKLYNTGGPGHWVTA*

Sequence 2527

Contig_0768_pos_5390_4893,
 10 is similar to (with p-value 3.0e-44)
 >gp:gp|U96108|SCU96108_7 Staphylococcus carnosus (3R)-hydrox
 ymyristoyl acyl carrier protein dehydrase homolog (fabZ) gen
 e, partial cds, YwpF homolog, single-strand binding protein
 homolog (ssb), SceD precursor (sceD), SceA precursor (sceA)
 15 and SceE precursor (sceE) genes, complete cds, and TenA homo
 log (tenA) gene, partial cds. NID: g2735509.
 atggaagatgttaaatttctagtagaacaattgagtttatgttagaaggagaagttgaa
 gcgcgatgaggttctagctgatttcattaatgaaccatatgaagaaatagtaaaagaaaaa
 gtatggccaccaagtggatgattatataaacatatgtacttcaatgcatttgacgt
 20 gaaaatgcagccttcacgattgcagcgatggcaccctgtccatacgtctacgctgtcatt
 ggtaaacgtgcgatggaagatcccaaattaaataaagaatcagtgacttctaaatggttt
 caattttatagtagtgaatggacgaacttgttgatgtgttcgatcaattgatggaccgt
 ttaactaaacattgtagttagacagaaaaaagagattaaagaaaatttcttgcaaagt
 actattcatgagagacatttcttcaatatggcatatattaatgaaaaatgggaatatggg
 25 ggaataacaatgaataa

Sequence 2528

MEDVKFLVEQIEFMLEGEVEAHEVLADFINEPYEEIVKEKVWPPSGDHYIKHMYFNAFAR
 ENAAFTIAAMAPCPYVYAVIGKRAMEDPKLNKESVTSKWQFYSTEMDELVDVFDQLMDR
 30 LTKHCSETEKKEIKENFLQSTIHERHFFNMAYINEKWEYGGNNNE*

Sequence 2529

Contig_0768_pos_4639_4079,
 is similar to (with p-value 8.0e-41)
 35 >sp:sp|P44697|THID_HAEIN PHOSPHOMETHYLPYRIMIDINE KINASE (EC
 2.7.4.7) (HMP-PHOSPHATE KINASE) (HMP-P KINASE). >pir:pir|I64
 151|I64151 hypothetical protein HI0416 - Haemophilus influen
 zae (strain Rd KW20) >gp:gp|U32725|U32725_3 Haemophilus infl
 uenzae Rd section 40 of 163 of the complete genome. NID: g15
 40 73387.
 atggagtttgattcaaagttactttaaaggaacattccaatattccttatgtgattgatcca
 gttatgcttgccaaaagtggtgattcattgatggatgaaaataactaaaaatcatttgcaa
 tcacattattacctttggctgatgttgctactccgaatatacctgaagctgaggaaatt
 actggtattaaaaatgaatgatgaagaaagcatagcgtaaagcaggtcaaatctttattaat
 45 gaaattggtagtaaggggtgttgtaattaagggagggcattcagccgatttaataatgct
 aaagattttcttttactaagaatgaaacgtatacctttgagaacaaacgctttgatact
 aaacatactcatggaactggttgactttttcagcagttattacagcagaattagctaaa
 ggtcgttccataaaaagatgcagttaaaaaagcaaaagagtttatttcattaagattgaa
 cataccccagaaattggcaaggggagaggacctgtaaatcattttgcttatatgaagaaa
 50 gtaggttttagatgatgaataa

Sequence 2530

MELIQSYLKEHSNIPYVIDPVMLAKSGDSLMDENTKNHLQSTLLPLADVTPNIPAEAEI
 TGIKINDEESIRKAGQIFINEIGSKGVVIGKGHSA DLNNAKDFLFTKNETYTFENKRFD
 55 KHTHGTGCTFSAVITAELAKGRS IKDAVKKAKEFISLSIEHTPEIGKGRGPVNHFA YMKK
 VGLDDE*

Sequence 2531

Contig_0769_pos_1300_1833,

is similar to (with p-value 1.0e-69)
 >sp:sp|P37948|GLPT_BACSU GLYCEROL-3-PHOSPHATE TRANSPORTER (G-3-P TRANSPORTER) (G-3-P PERMEASE). >pir:pir|S37250|S37250 glycerol-3-phosphate transport protein - Bacillus subtilis >g
 5 p:gp|Z26522|BSGLPTQ_1 B.subtilis glpT and glpQ genes for glycerol 3-phosphate permease and glycerophosphoryl diester phosphodiesterase. NID: g403371. >gp:gp|Z99105|BSUB0002_43 Bacillus subtilis complete genome (section 2 of 21): from 194651 to 415810. NID: g2632457. >gp:gp|AB006424|AB006424_44 Bacillus subtilis genomic DNA, 70 kb region between 17 and 23 degree. NID: g3599592.
 10 atgtttaatttcttaaaaccagctagacatatcaaatcattgccatcagaaaaagtagatgatacgtataaaagactacgctttcaagtcttttaggaatatttatcggttatgctggttactatttattaagaagaacttctcttttagcaatgccttcattaattgagcaaggcttt
 15 agtataagggaattaggtattgcattatctgcagtatctatcgcatatggcttttagcaaa
 tttgtaattgggcactgtcagcgatcgaagtaatgctcggtatgttcttaacttttaggttta
 gtattgacagcaattattaacttattatttaggatttattccattctttacttcaagcata
 actattatgtttatcatgctgttttttagttggatggtccaaggaatgggctggccacca
 tctggacgtgtgttagttcattgggttagtggtcagcgaaacgtggaagcaaacgtcaata
 20 tggaatgtagcacataatgtaggcggaggtttaatggcacctattgtctacgtag

Sequence 2532
 MFNFKPARHIKSLPSEKVDITYKRLRFQVFLGIFIGYAGYYLLRKNFSLAMPSLIEQGF
 SKGELGIALSAVSIAYGFSKFVMGTVSDRSNARMFLTLGLVLTAIINLLGFIPFFTSSI
 25 TIMFIMLEFLVGWFGMGWPPSGRVLVHWFVSVSESGSKTSIWNVAHNVGGLMAPIAT*

Sequence 2533
 Contig_0769_pos_1843_2658,
 is similar to (with p-value 2.0e-72)
 30 >sp:sp|P37948|GLPT_BACSU GLYCEROL-3-PHOSPHATE TRANSPORTER (G-3-P TRANSPORTER) (G-3-P PERMEASE). >pir:pir|S37250|S37250 glycerol-3-phosphate transport protein - Bacillus subtilis >g
 p:gp|Z26522|BSGLPTQ_1 B.subtilis glpT and glpQ genes for glycerol 3-phosphate permease and glycerophosphoryl diester phosphodiesterase. NID: g403371. >gp:gp|Z99105|BSUB0002_43 Bacillus subtilis complete genome (section 2 of 21): from 194651 to 415810. NID: g2632457. >gp:gp|AB006424|AB006424_44 Bacillus subtilis genomic DNA, 70 kb region between 17 and 23 degree. NID: g3599592.
 35 atgactgcattatacaacttcggttatattaaagggtttgaaggtgtctttatataccctgcactattggttatcattattgccatcttgtcttacatactaattagagatacaccacaa
 tctcagggtttaccaccaattgagcagtatataaaatgattatgccacttcaactaaacaa
 acaattgaaacagaactaactacgaaagaaatattatttaaataatgtacttaataacaaa
 tgggtatgggcgattgcttttacaacattttcgtttattttgtgcgttatgggttttg
 40 gactgggctccgacatacttaagttaggaaaagcattttgatttaagtgcctcaggttgg
 gcttacttcttatacgaatgggcaggaattccgggcacgctactctgtggttatctatct
 gacaaactatttaaaggctgcgcgtggtccagcaggcttcttctttatgttaggcgtaaca
 atctttatccttatataattgggttaaatccaccaggtcacgcatgggttagataatctttca
 ttaattgggtattggtttcttaatttatggtccagtcagttaattggtttacaagcgtaa
 45 gattatgttcctaaaaagcagctggcactgcggctggtctaactggtctatttggtatc
 ttattcggtgcagttatggctaataattgtattagggtttgtagttcaacattttggatgg
 catattggctttgtgttattaacagtcacgcatactcgctatgttatgtttcatttta
 acttggataaacgtggtcaagaacaaatcgactag

Sequence 2534
 MTALYNFGYLKGFEGVFIYPALLVIIIAILSYILIRDTFQSQGLPPIEQYKNDYATSTKQ
 TIETELTTKEILFKYVLNNKWVWAI AFTNIFVYFVRYGVLDWAPTYLSEEKHFDSLASGW
 AYFLYEWAGIPGTL LCGYLSDKLFGKRRGPAGFFFM LGVTIFILIYWLNP PGHAWLDNLS
 55 LIGIGFLIYGPVMLIGLQALDYVPKKAAGTAAGLTGLFGYLF GAVMANIVLGFV VQHFVGW

HIGFVLLTVISILAMLCFILTWNKRGEQID*

Sequence 2535

Contig_0769_pos_4799_5716,

5 putative peptide of unknown function

atggttatttcaaagaaattagcatttatcactataggtatcttaatatgtattctttta
 agttcacctctgatcgatcaacagatttcaaattatttcatgaatcaagattcaaatttc
 ggtactttatttcaaaactacggtttatttccacctacgctcatattaattatttcaact
 gtaattttaaattattatattttaaactacatttcaaaataaattagcaaaaacgttaact
 10 ttgattttatcttttatctttacggttgataaaaaaccaatgaattcgtgtcagaaaccgct
 caatatatgttatctacatcagaaaaacatcaaaaaatcataagccaatgggtatggcaaat
 aatgaaggaaacgctggaaatgctttatcttttaggaatgagttttttatttctttgatt
 atcattataattatcactattatttggttatcagttttgggttaaaacatacaaaataacca
 gaacttgatcatttattcaaagtttcacttattagctttatgattttatttattggatta
 15 gaactcgttgatagcttgaaacatttatggggacggttttagaccatatgaaatcactgac
 aaagctggacatttcaactcactggcttacaataaatggaaacactggacatagttccttc
 ccttcaggacacactggaaatggtgcatttctaattgtttatcgcatcttactttaaaaa
 ttacgcactcaaaaaatagtggttagtatcggtatgttatagtttttaaatggcattg
 agcagaattcgtattgggtgcacatttcaactagtgatgaacaatgagcttgcttattatg
 20 ttttcaactcatggttattgcagattttatcattaatcaagtcatatctatacataaaaaac
 aagctaagtaagtcctag

Sequence 2536

MVISKKLAFITIGILICILLSSPLIDQQISNYFMNQDSIFGTLFQNYGLFPPTLILIIIST
 25 VILNYYILTTFQNKLAKTLLILSFI FTLIKTNFVSETAQYMLSTSENIKNHKPMGMAN
 NEGNAGNALSLGMSFFISLIIIIITIIICYQFWLKHNTNNQELDLHLFKVSLISFMILETGL
 ELVDSLKHLWGRFRPYEITDKAGHFTHWLTINGNTGHSSFPSGHTGNGAFLMFIA:YFKK
 LRTQKIVFSIGLCYSILMALSRIRIGAHFTSDVTMSLLIMFSLMVIADFIINQVISIHK
 N
 KLSKV*

30

Sequence 2537

Contig_0769_pos_7866_7003,

putative peptide of unknown function

atgccgttaactcttaggtatcgaaaaagggttggtttaagaagaatcattagaaattgag
 35 atgattgaaccaaaggaacactttgacgcactagacgagattgaaaatgggttcaatggat
 atcgcgattactgaacctattcatctggttcaagatagagctaaagaacaaaaagtcac
 ggggtttgcgagatatcttcacactaatggaggtattatgtataacaagataaaaaatc
 gctcgcccaagatttaacggttaaacggcttcaatatcccggtgctcccggtccaggt
 ggtattgctatggctaaaacgatgattgaagctgatgggtgtacatttgaagaaggtgac
 40 atcacaccagttaatcatagttttatcatactgatgcacttttaactgataaagctgat
 gctgctacactcattttcgaaaattttgaaattcttgaagctagaaatcaaggactta
 gtagattattttccacttaaaaattataatgtacccgatttttgcactcattttcatt
 acaacacctgaagtattaaatcagagtgaaggttaaactcaaaaagttcctaaaaatcatt
 caaaaaacaattcattacatcaatcatcatttagaaagtgaattgaaatttactcaacg
 45 tatactcaaaccgatattcttaaccaattaaacaaagatacaattgaagcaacagctaaa
 tgttttacaatatgatttgcctatgagttccgactactacaatgatttacagttgtggctt
 aaagaagtcaatgatattaaagacacgattaaccaactttatattttacaatatcaacta
 ctattcagtcgatatacaaaaataa

50 Sequence 2538

MPLILGIEKGWFKESLEIEMIEPKEHFDALDEIENGSMEDIAITEPIHLVQDRAKEQKVI
 GFARYLHTNGGIMYNKDKNIARPKDLIGKRLQYPGAPGGIAMAKTMIADGGTFFEEGD
 ITPVNHSEFYHTDALLTDKADAATLIFENFEILEARNQGLNVDYFPLKNYNVPDFCQLIFI
 TTPEVLNTSEVKLKKFLKIIQKTIHYINHHLESAIEIYSTYTQTDISNQLNKDTIEATAK
 55 CFTNDLSMSSDYNDLQLWLKEVNNDIKDTINPTLYFTNQLLFSRYTK*

Sequence 2539

Contig_0769_pos_6570_6241,

putative peptide of unknown function

atggggtttgaagagaaattggaggttgagaaagatgtatgtagagagaaagcctgcactt
 tatatcgaagatttacgagacgaattcaaaaatagcttaaaacattttaagatgatgat
 gaagcatttaatacatttagttggatttgggaactggaccacctatattcttcagcactt
 aaagaaataagtacaaaattaagtatttttagatgataactttaactatcaatataaacac
 5 aatccaattcatcatatggaacgacgagtaaaagagatgcacagtttagttaaaaagttag
 aatcgtaaaggatttgaagtaagtgcataa

Sequence 2540

MGLKRNWRLRKMYVERKPALYIEDLRDEFKNSLKHFKDDDEAFNTLVGFVELDHLHYSSAL
 10 KEISTKLSILDDNFNYQYKHNPIHHMERRVKEMHSLVKKLNRKGFEVSA*

Sequence 2541

Contig_0769_pos_4499_3618,
 putative peptide of unknown function

15 atgtatatgattaagcaaaattttacaactaaagtttatccacttaacaatttgagggga
 gaatatagtgaagttgcctacgcaggagatagccaagctgatgatgtgattgtatttatt
 catggtgcacttttaacatataaaaattatgacgatgttcgaaccttacttcagagaaat
 aaatcaaattttattaattgtccaagtcgtggtagaagttcagatttagatcgtgcacaca
 20 catacattagattacgctgcacgtatataatgatgtattaacgcaaattgttaaggag
 caacaaataaaagaactgagttattgtcggttattcaatgggaggaatgattgcgacacgg
 ttacttaagtataatacattaccagtcctctcatcttattttattacatagcgcagcgaaa
 attactccagatgcaagtatgttagcacgattattcactagtgcagagtaagagagcagtg
 ttaaaagatgaaattaaggcagtgaaaaatcttcctcaatatatactagataaaacgatt
 25 tatgcacaaaaggaaaacgcacttgatttggtagaatttattgcaactattaaaactata
 attacggatatgatttacacgatttaatacagattatttaccagatatcgatgagattaaa
 caatttccgaaaatattatttattgtctggaaaagaagatcaaattattccttatacggat
 tctcaagctacgttagaaaagtttaaggcgttaggtggagaaactaaagaagttatttat
 ccaggaattggatcatatcgatttcccaagtgttttagaaaccaatcgatggacaaact
 30 ggcgtggtagatgaaattaaagcgtggatttcaaaaaataa

Sequence 2542

MYMIKQNFHTTKVYPLKQFEGEYSEVAYAGDSQADDVIVFIHGALLTYKIMTMFEPYFRDY
 KLIFINCPSRGRSSDLDRDTHLDDYAARIYDVLTQIVKEQQIKELSIVGYSMGGMIATR
 LLKYNTLPVSHLIYLHSAKITPDASMLARLFTSESKRAVLKDEIKAVKNLPQYILDKTI
 35 YAQKENALDLVQFIAPIKTIITDMIYTINTDYLDPIDIEIKQFPKILFMSGKEDQIIPYTD
 SQATLEKFKALGGETKEVIYPGIGHIDFSPVLETQSDGQTGVVDEIKAWISKK*

Sequence 2543

Contig_0769_pos_3394_2816,
 putative peptide of unknown function

40 gtggagaaacttgaactgattttactacttaaggaaataggttgacaaattaaagaaatt
 aaagtgttattaaaagatgattcatcaatgaaatctctttacacgtttcttcaaatcaaa
 aagcatgaaatacaacagtcataacagataaagaaaacaaggatgtaaaatagagcaa
 attcaacgctatgttcatcaaaattcaatttctccaattcattatttacaagatatagct
 45 atttatatggaagaatctcatagactcaaagggtgtcagaaagaaattatggttaagtata
 gctctgattggttcattacaatatggtggtttgatcacctcaattgtaactcaaaggaaa
 aagcctttcttgcgatgatgcctgtggttagctatgtattcactttggttaacgaaaaag
 tataaaaaagaatgtttcatatgtctgtcctaattgccaccatgtgtttaatcctagtgtc
 attcattttgtaacagcatcacatacacctaaaacaagaaagcttcaatgtcctgattgt
 50 caggaatgcattactgtgtagaaattgctaagctatga

Sequence 2544

VEKLELILLKKEIGCTIKEIKVLLKDDSSMKSlyTFLQIKKHEIQQSITDKENKVCKIEQ
 IQRYVHQNSISPIHYLQDIAIYMEESHRLKGVKKLWLSIALIGSLQYGLITSIVTQRK
 55 KPFLSMMPVVAMYSWLTKKYKKNVSVCPNCHHVFNPSVIHFVTASHTPKTRKLCPCD
 HEMHYCVEIAKL*

Sequence 2545

Contig_0769_pos_850_251,

is similar to (with p-value 1.0e-51)
 >sp:sp|P10585|GNTR_BACSU GLUCONATE OPERON TRANSCRIPTIONAL RE
 PRESSOR (P28 PROTEIN). >pir:pir|C26190|C26190 transcriptional
 1 repressor GntR - Bacillus subtilis >gp:gp|AB005554|AB00555
 5 4_4 Bacillus subtilis genomic DNA, 36 kb region between gnt
 and iol operons. NID: g2280496. >gp:gp|J02584|BACGNT_1 B.sub
 tilis (gluconate operon) gntR, gntK and gntP genes encoding
 gnt repressor, gluconate kinase and permease, and gntZ gene.
 NID: g143013. >gp:gp|Z99124|BSUB0021_110 Bacillus subtilis
 10 complete genome (section 21 of 21): from 3999281 to 4214814.
 NID: g2636442.
 gtgagtggtgagattgaatcccatatataactaagtgaataatcaagttgcacaacaattt
 aatgttagtcggtcaccagtcagagatgcatttaagttacttcaaacagatcaactgatt
 caattagaacgtatggcgctcaggtacttcttttgggtgatcaagagaaaaaggaaatg
 15 tatgatttcggtttgatgctcgaatcattcgcttttcaaaattaagtggaacagataca
 caacatattattaaggaaatgaaaagcaattagaaatgatgaagggtgcagtcacaattt
 gaggatgctgaagcatttacacaacatgattttgagtttcatgaggtgatgattcaagca
 acaaacatcagtatcttaaaagtgttttggaatcaccttaaacctgttatggaatcactc
 atactcatttcaatgagacaaagaatggcaaatgaccccaaagatttcgagagaattcat
 20 aaaaatcatcaagtttttatagatgctgttgagaacgatgatgcctccatattgagaaaa
 gcattccatttaaattttgatgatgtaggagaaaatattgaagcattctggttacgttaa

Sequence 2546

25 VSGEIESHIQLSENQVAQQFNVSRSFVRDAFKLLQTDQLIQLERMGAQVLPFGDQEKEM
 YDLRLMLESFAFSKLSGTDQHIKEMKKQLEMMKVAVQFEDAEAFQHDFEFHEVMIQA
 TNHQYLKVFWNHLKPVMESLILISMRQRMANDPKDFERIHKNHQVFIDAVENDDASILRK
 AFHLNFDDVGENIEAFWLR*

Sequence 2547

30 Contig_0770_pos_3341_3688,
 is similar to (with p-value 3.0e-41)
 >sp:sp|P37527|YAAD_BACSU 31.6 KD GUANYLYLATED PROTEIN IN DAC
 A-SERS INTERGENIC REGION. >gp:gp|D26185|BAC180K_75 B. subtil
 35 is DNA, 180 kilobase region of replication origin. NID: g467
 326. >gp:gp|Z99104|BSUB0001_11 Bacillus subtilis complete ge
 nome (section 1 of 21): from 1 to 213080. NID: g2632267.
 atgtctaaaatagtaggatcagatcgagttaaaagaggaaatggctgaaatgcaaaaaggc
 gggtgctattatggacgtcggttaatgcagaacaagctaaaattgctgaagaagccggagct
 40 gttgccgtaatggcattagagcgtgtaccatcagatattcgctgctgctggcggtgttgca
 cgtatggcgaatcctaaaatagttgaagaagttatgaatgccgtatcaattccggttatg
 gctaaagccagaattggtcatattacagaagctagagttttagaatcgatgggtacacgg
 tttcaagaacccaacatactacaaacgaatttcaaaaggcgagagtaa

Sequence 2548

45 MSKIVGSDRVKRGMAEMQKGGVIMDVVNAEQAKIAEEAGAVAVMALERVPSDIRAAGGVA
 RMANPKIVEEVMNAVSI PVMAKARIGHITEARVLESMGTRFQEPNLTNFKRRE*

Sequence 2549

50 Contig_0770_pos_5637_5939,
 is similar to (with p-value 3.0e-42)
 >gp:gp|AF016634|AF016634_1 Lactococcus lactis cremoris ClpB
 chaperone homolog (clpB) and phosphoribosylformylglycinamide
 cyclo-ligase (pur5) genes, complete cds; and phosphoribosyl
 55 glycinamide formyltransferase (pur3) gene, partial cds. NID:
 g3150045.
 atgattcgaccttcagattcttttacttctttcaatactgcttttaaacgttcttcaaat
 tcacctctatattttgacactgcaactaaggcacttaaatcaagctcgaaaatcgtttta
 tcgagtaatgattctggaacgtctttactgtacaattcggtgtgctaaaccttcaacaatt

gcagctttiacctacacctgggttcaccgattaaaaccggattatTTTTTgTTTTTcgaactt
aatatacgaattgtattacgtatttcttcatctctaccgatgacaggggccattttacct
tga

5 Sequence 2550

MIRPSDSFTSFNTAFKRSSNSPLYFAPATKALKSSSKIIVLSSNDSGTSLRTIRCAKPSTI
AVLPTPGSPIKTGLFFVFRNLNIRIVLRISSSLPMTGSILP*

Sequence 2551

- 10 Contig_0770_pos_4797_3853,
is similar to (with p-value 0.0e+00)
>gp:gp|AF016634|AF016634_1 Lactococcus lactis cremoris ClpB
chaperone homolog (clpB) and phosphoribosylformylglycinamide
cyclo-ligase (pur5) genes, complete cds; and phosphoribosyl
15 glycinamide formyltransferase (pur3) gene, partial cds. NID:
g3150045.

- gtgaaacagaaagagaaaagctattaagttaaagtgcacatcttacacaaacgtgtagta
ggtcaagataaagcagttgatttagtatcagacgcagtagtttagagcacgtgctggaatt
aaagatccgaatagaccaatcggaagtttcttattcttaggacctactggagtaggtaaa
20 actgaattagcaaaaatcgcttgcttcacacttttcgattctgaaaaacatatgattaga
attgatatgagcgaatatatggaaaaacatgctgtatcacgtttaattggcgacactcca
ggttatgtaggtcacgatgaaggtggtcaattaactgaagcagtttagacgtaaatccatac
tcagttattttgtagacgaagttgaaaaagcacatagcgatgttttaattgtattactt
caaatactagatgaaggtcgctctacggattctaaaggtagaagtggtgactttaaaaat
25 accattatcatcatgactagtaaatattggttcacaagtattacttgaaaatgtaaaagat
gctggtgaaattagtgatgatacagagaaagcagttatggacagtcctacatgcatacttc
aaacctgaaatattaaatcgctatggatgacatcggttatttaaacattatcagttaat
gatatgagtagtattgtagataaaaattttaacacaattaaatatgagattattagatcaa
catatctcaattgaagtgcagagaagaagcgaaaaaatggctaggtgaagaagcgtatgaa
30 ccacaatttggtgcaagaccattaaaacgctttgttcaacgacaaatagaaactccaatt
gcacgtatgatgattaaagaaagtctacctgaaggtacaataattaaagtagatttaaat
gacaataaagaacttgattttaagggtgttaaactacgtcttaa

Sequence 2552

- 35 VETEREKLLSLSDILHKRVVQDKAVDLVSDAVVRARAGIKDPNRPISFLFLGPTGVGK
TELAKSCLASSLFDSEKHMIRIDMSEYMEKHAVSRILIGAPPGYVGHDEGGQLTEAVRRNPY
SVILLDEVEKAHSDVFNVLQILDEGRITDSKGRSVDFKNTIIIMTSNIGSQVLENVKD
AGEISDDTEKAVMDSLHAYFKPEILNRMDDIVLFLKPLSVNDMSMIVDKILTQLNMRLLDQ
HISIEVTEEAKKWLGEAYEPQFGARPLKRFVQRQIETPIARMMIKESLPEGTIKVDLN
40 DNKELDFKVVKPTS*

Sequence 2553

- Contig_0770_pos_1762_383,
is similar to (with p-value 0.0e+00)
45 >sp:sp|P39616|DHA2_BACSU PROBABLE ALDEHYDE DEHYDROGENASE YWD
H (EC 1.2.1.3). >pir:pir|S39713|S39713 hypothetical protein
- Bacillus subtilis >gp:gp|X73124|BSGENR_59 B.subtilis genom
ic region (325 to 333). NID: g413923. >gp:gp|Z99123|BSUB0020
_91 Bacillus subtilis complete genome (section 20 of 21): fr
50 om 3798401 to 4010550. NID: g2636240.

- atgacaataattagagataaatttaacaatagtaaagcttttttaatacgcataaaaca
aaaaaccttaaatctcgaaaaacaacaacttaattactaaagtaaaaaatataaaaaatcat
gaaaatgaattattagatgccttatataaagatttagttaaagtaaggtgaagcatatc
gcaactgaaattggtatgcttttgaaaaagcataaagctaatgcgcaaaagagttaaaaaat
55 tggtcgaaaaccaaacaacggatacaccactctacttattccctacaaagagttatatt
aaaaaagaaccttacggtacggtgcttattataggaccatttaattatccggttcaatta
gttttcgagcctctcatcgagcaatagctgccggaataactgctatagttaaaccttca
gagttaacacctcatgttgccattgtgatcaaggacatcattgaagatacatttgatgaa
gcatacgtttctgttggtgaaggtggtattgaagaaacccaaacggttattaagtctcca

tttgattatatgttctttactggcagtgaaaaagtcggaaaaattgtctatgaagctgca
 gcaagaaaattaattccagttactcttgaacttggcggtaaatcacctgtcattgtcgat
 gatacagccaatatcaaagtagcgagtgaaactgattagttttggtaaatctactaatgct
 5 ggtcaaacatgtgtcgctccagattatatattagttcagcggaaagttaaaaaatgattta
 ataaaaagctcttaaaaaacaattactgaattttacggagaaaaatattgaaaaagccct
 gatttcggacggattgttaatacaaaaacactttaatcgggtgaatgacttgattcaaat
 cataaagataatgttgtttttggaggtaatagttctaaagaagatttatattgaacct
 actttattggataacataaccaatgacaataaaatcatgaaagaagaatattcgggtccc
 attttgctattattacttatgataatttcgatgaagtacttgaaatcatccaaagtaaa
 10 tcaaaaccactaagttgtatcttttttagcgaagatgaaaacatgacacatagagtgggt
 gaagaattatcatttggggcggtgcaattaacgatacgttaatgcatttagctaactct
 aacttaccttccgggtgtaggttcttcaggcataggtcaatatcatggtaagtattct
 tttgatacatttagtcataatgaaatcacacatttaaatctacacgtctagaatcgagt
 ttattttccctccatataaaggtaaatttaatatattaaaaccttcttcaagaactag
 15

Sequence 2554

MTIIRDKFNNKAFNTHKTKNLKFRKQQLKLLSKNIKNHENELLDALYKDLGKSKVEAY
 ATEIGMLLSIKLMRKELKNWSKTKQTDTPLYLFPTKSYIKKEPYGTVLIIGPFNYPVQL
 20 VFEPLIGAIAGNTAIVKPELTPHVAIVIKDIEDTFDEAYVSVVEGGIEETQTLLSLP
 FDMFFTGSEKVGKIVYEAARKLIPVTLELGGKSPVIVDDTANIKVASERISFGKFTNA
 GQTCVAPDYILVQRKVKNDLIKALKKTITEFYGENIEKSPDFGRIVNQKHFNRLNDLIQI
 HKDNVVFSGNKKEDLYIEPTLLDNITNDNKIMKEEIFGPILPIITYDNFDEVLEIIQSK
 SKPLSLYLFSEDENMTHRVVEELSFGGGAINDTIMHLANPNLPFGGVGSSGIGQYHGKYS
 25 FDTFSHMKSYTFKSTRLESSLFFPPYKGFYIKTFFKN*

Sequence 2555

Contig_0771_pos_3650_4303,
 is similar to (with p-value 7.0e-69)
 30 >gp:gp|AF008219|AF008219_3 *Borrelia afzelii* R-IP3 chromosome
 right end, arcA and arcB genes, complete cds. NID: g2697111

atgaaaaatttacgtaacagaagcttttttaactttattagacttttcacgacaagaggta
 gaatttttattaacactctccgaagatttgaagcgtgccaaatatatcggcactgaaaag
 35 cctatgctaaaaataaaaaatatcgcgcttcttttgaaaaagattccactagaacacgt
 tgcgcattcgaagttgccgcacatgatcaaggtgcacacgtcacttatcttggacctaca
 ggttctcaaatgggtaaaaaagaaactgctaaagatacagcacgtgtacttggtggtatg
 tatgatgggtattgagtaccgaggtttctctcaacgtactgtagaacattagcgcaatat
 tcagggtgtccggtatggaatggattaaccgatgaagatcacctacacaagtgttgc
 40 gactttttaactgctaaagaagtattgaaaaagagtatgctgatatcaactttacttat
 gttggcgatggacgtaacaatgttgctaacgcattaatgcaaggtgctgccattatgggt
 atgaatttccatcttgtttgtcctaaagaactcaatccgacagaagaattattaaatcgt
 tgcgaacgtattgcgacggaaaatggcggtaacattttaatatcattacttta

Sequence 2556

MKNLNRNSFLTLDFSrqEVEFLTLSEDLKRAKYIGTEKPLKNKNIALLFKEDSTRTR
 CAFEVAHDQGAHVTYLGPTGSQMGKETAKDTARVLGGMYDGIEYRGFSQRTVETLAQY
 50 SGVPVWNGLTDEDHPTQVLADFLTAKEVLKKEYADINFTYVGDRNNVANALMQGAAIMG
 MNFHLVCPKELNPTEELLNRCERIATENGGNILISLL*

Sequence 2557

Contig_0771_pos_5383_4310,
 putative peptide of unknown function
 atgattaattcattggaatggttgtagtaatttggtagtggttttttgtttgataaa
 55 cttggtggttacaaaacgatttttaataaggtacattttacgtgtttatgtagtaccacatta
 ctcaacttgtttcatggctggccatggtatgcaatttggtagtactacttggtttggt
 ggcggaatgatagttcctgctattttatgcgatggcaggtgccgtttggcctaattggagga
 agacaaacttttaagtctatctacttagcacaataataggggttgcttttaggcgcggca
 ttaggaggttttggctgaatttagtttcaattatatttttatggcaaacctcattatg

tatgtactttttgccatcggtggcgattacacagttcaatttagagattaatgctaaattt
 aaaccacaagattcgatagatttaaaaagcaaaagaaaaataaaaaacgatttactgctatg
 atgctagtagtgcaaatgtttgcaatttgttggttgcatatattcaatgggaacaacg
 atagcttcattcacacaatcaattaatatttcaatgtctcaatagtgattatgaca
 5 attaatggaattatgatttttagtagctcaacctttaataagaccaattattatcttatta
 aaaggttaatttaaacatcaaatgtttgtaggtattttaattttatgagttctttccta
 gtgacgagttttgcaaatcactttgctatattttagtggcatggtcattttaactttt
 ggagaaatgtttgtttggcctgcagtaccaactatagcaaatcaacttgcaccagttgga
 aagcaagggcaataccaaggtttgttaattcagcatctacagtgggtaaagcatttgga
 10 ccatttattgggggtatactttagatatactttaatatgagtagtgatttattgggatg
 attataattattaagttttgcactgttattttaagtttctatgataaagtgttaccgaag
 aattttaaaaatcaacatcaatcaagaagacgacgaaatcagaatggtatttaa

Sequence 2558

15 MINSEGMVVGNNLLGGFLFDKLGKYKTIILIGTFTCLCSTTLLNLFHGWPHYAIWLVLGGFG
 GGMIVPAIYAMAGAVWPNGGRQTFNAIYLAQNIGVALGAALGGFVAEFSFYIFMANLIM
 YVLFIVAIVITQFNLEINAKFKPQDSIDLKSKENKKRFTAMMLVCAMFAICWIAIYQWETT
 IASFTQSNISMYSQYSVLWTLINGIMILVAQPLIRPIIILLKGNLKHQMFVGILIFMSSFL
 VTSFANHFIAIFVVGMLTTFGEMFVWPVPTIANQLAPVGKQGYQGFVNSASTVKGAFG
 20 PFIGGILVDTFNMSMMFIGMIILLSFALLFLSFYDKVLPKNFKNQHQSRRRRNQNGI*

Sequence 2559

Contig_0771_pos_2844_2416,
 is similar to (with p-value 2.0e-50)
 25 >pir:pir|S58181|S58181 fofB protein - Staphylococcus sp. >gp
 :gp|X89875|SSPPDNAFB 1 Staphylococcus sp. plasmidic DNA for
 fosB gene. NID: g927563.
 atggaaataacaaatgttaatcatatttgtttttcagtgagtgatttaaacctctata
 caattttataaagatattttacatgggtgacttattagtagatcagatagaacgacagcatat
 30 ttaactattggtcatacttggattgcactgaatctagaaaaaaatataccaaggaatgaa
 ataagtcattcctatacgcacgttgctttctccatagatgaagaagattttcaacagtggt
 attcaatggcttaaaagagaatcaagtaaatattttaaaagggcgaccaagagacattaaa
 gacaaaaaatcgatataattttacagatctggatgggcataaaattgaattacatactgga
 acattaaaaagatagaatggaatattataaatgtgagaagacgcatatgcaattttacgat
 35 gagttttga

Sequence 2560

MEITNVNHICFSVSDLNLSIQFYKDILHGDLLVSDRRTAYLTIGHTWIALNLEKNIPRNE
 ISHSYTHVAFSIDEEDFQQWIQWLKENQVNILKGRPRDIKDKKSIYFTDLGDGHKIELHTG
 40 TLKDRMEYYKCEKTHMQFYDEF*

Sequence 2561

Contig_0771_pos_1176_178,
 is similar to (with p-value 0.0e+00)
 45 >sp:sp|P53557|BIOB_BACSU BIOTIN SYNTHETASE (EC 2.8.1.-). >gp
 :gp|AF008220|AF008220 77 Bacillus subtilis rrnB-dnaB genomic
 region. NID: g2293135. >gp:gp|U51868|BSU51868_5 Bacillus su
 btilis biotin biosynthetic operon genes, complete and partial
 1 cds. NID: g1277024. >gp:gp|Z99119|BSUB0016_93 Bacillus sub
 50 tilis complete genome (section 16 of 21): from 2997771 to 32
 13410. NID: g2635411.
 atgctaatttttaagaaaaaggagttaaagattatgacattaaacctagctcaacgtgtg
 ttaaatcaagagtcatttaacaaaagatgaagcaatatctattttcgaaaatgctgaaatt
 gatacatttgattttatataatgaagcctacacagtgagaaaacattactatggtaaaaaa
 55 gtttaagcttaatatgatattaaatgctaaaagtggatctgtgcagaaaattgtgggtac
 tgtgggcaatctgtaaaaatgaaagaaaagcaacgttatgcacttgttgaaacaggaccaa
 attaaagaaggcgctcaagtggcaactgaaaatcaaatcggtacatactgtattgttatg
 agtggttagaggtcctagtaacagagaagtcgatcatatttgcgaaacagtagaagatatt
 aaaaagatacacccacaactaaagatttgcgtgcttaggattaacgaaagaagaacag

gctaaaaaattaaaggctgctggtgtcgatcggttataatcataatttaaacgagtgag
 cggtatcacgatgaagtagtaactacacatacatatgaggatagagtgaatacggttgaa
 atgatgaaagataataatatttctccttggtcaggtgtgatatgtggtatgggagagtcg
 aatgaggacattattgatatggcatttgctttaagagccatcgatgctgatagcattcct
 5 attaattttttacatcctattaaaggaactaaatttggtggattagatttattgtcacca
 atgaaatgtttaagaattatagcgatgttttaggttaataatccaacaaaagaattcga
 attgcaggtggacgggaggttaaattctacgttcattacaaccactcgattgaaagcggct
 aattcaattttttaggagattacttaattacaggcgggtcaaccgaatgaggaagattat
 cgcatgattgaagatttaggggttgaaatcgacagttaa

10

Sequence 2562

MLIFKKKELKIMTLNLAQRVLNQESLTKDEAISIFENAEIDTFDLLNEAYTVRKHYGKK
 VKLNMILNAKSGICAENCGYCGQSVKMKKQRYALVEQDQIKEGAQVATENQIGTYCIVM
 SGRGPSNREVDHICETVEDIKKIHPQLKICACLGTLKEEQAKKLKAAGVDYRNHNLNTSE
 15 RYHDEVVTTHTYEDRVNTVEMMKDNNISPCSGVICMGESNEDIIDMAFALRAIDADSIP
 INFLHPIKGTGKFGGLDLLSPMKCLRIIAMFRLINPTKEIRIAGGREVNLRSLQPLALKAA
 NSIFVG DYLTGGQPNEEDYRMIEDLGFEIDS*

Sequence 2563

20 Contig_0774_pos_436_873,
 is similar to (with p-value 3.0e-34)
 >sp:sp|P39149|UPP_BACSU URACIL PHOSPHORIBOSYLTRANSFERASE (EC
 2.4.2.9) (UMP PYROPHOSPHORYLASE) (UPRTASE). >pir:pir|S49364
 |S49364 uracil phosphoribosyltransferase - Bacillus subtilis
 25 >gp:gp|Z38002|BSSPORUPP_10 B.subtilis spoII-R, glyC and upp
 genes. NID: g556877. >gp:gp|Z99122|BSUB0019_186 Bacillus su
 btillis complete genome (section 19 of 21): from 3597091 to 3
 809700. NID: g2636029.
 atgttaaaagtattcaacactgaatcaagcatctctctatgttggcggtgactacaaca
 30 acgggttccaggtcagaatctttctctaacgttttaataacggagccattttatagct
 tcaggcctagttccaaatatggtcataactttttcatcaaaaatatctatctccgaatc
 tattattttagtaccgaataatctatccctgcatacctaaccctgggtgaatatacgtt
 ttgtcatttaatttttcatctaataatgccgaatataatctacatctgggtgtgcttct
 tgcattttttcaacgccttcagggcgagctattaacacataaaaacgtataacttttagct
 35 ccacgtttttttaatgaagaaattgcttcaatagctgaagcaccagtagcaagcaatagga
 tcaaccacaataatttga

Sequence 2564

MLKVFNTESSISLCWAVTTTGSRSSEFSNVLINGAIFIASGLVPMNVITFFIKNIYLRI
 40 YYLVPNNLSPASPKPGVIYALSFNFSNAAIYISTSGCASCIFSTPSGAAIKHIKRILLA
 PRFFNEEIASIAEAPVASIGSTTII*

Sequence 2565

45 Contig_0774_pos_2423_3718,
 is similar to (with p-value 9.0e-29)
 >pir:pir|S57509|S57509 integral membrane protein - Streptomy
 ces pristinaespiralis >gp:gp|X84072|SPDNAPTR_1 S.pristinaesp
 iralis ptr gene. NID: g872305.
 gtgccattacaatcatcatacaatagtgatattggtactattaatatagcagtttagctta
 50 tcggcactattttctggtctgtttattgttaggtgcaggagatattgcagataaaattggt
 agagtaaaagtgcgtacataggttagcacttaatatgttgggtcgattttaatcatt
 attacgccattaccaagtctattgattattggacgtgctattcaaggattgtcagcggca
 tgtataatgccagcgacactcgcaatcattaatgaatattatatcgggacagcacgacaa
 cgtgcattaagtactggtccatcggttcattggggaggtagtggtgtttgtactttgttt
 55 ggtggttttaattggcaactaaccttggtggtcgctcaatctttattgtttcaattaatctg
 acaatattatccatgtttctcattaaacatacacccgaaacaaaagcagagcctatagga
 gatcaaccgacagagacaaaagaatttgatgttgttgggttaataatcatcttagtggttagt
 atgttaagtattaatgtgataataactcaaacctctcaatttggtttgttctcaccattt
 attttgggacttattgcaatttttgttatatcggttaattatattcggtgatttacgaaaat

5 aaaatcaaaacaccacttgtcgattttgatatctttaaaaacaaaggttatacaggtgca
 acgatttcaaactttatgttgatggtgtagctggtgtacattgattgtagtaaatact
 ttttatcagcaaaaattagatttttaactctcaggaaacaggatataatttcacttacatac
 ctaattgcagtattaattatgatacgtgtgggtgaaatgatattacaatcgttaggacct
 10 aaaagacctttgttactcggaagtgccttgaccgtcataggattaatattattatctttg
 acgttttttacctaattgcttggtatatagcgtcaagtgtcattgggtattttattttggg
 accggttttaggtgtttatgcaacaccatccacggatacagctgttgcaaacgacragat
 gataaagtaggcgtcgcatccggtgtatataagatggcatcatccttaggaaatgcattt
 15 gggtgtggccatctcaagtacagtttacagcgtacttgcagcccaacttaatctgacttta
 gggtgttttactggagtaatgtttaatgcgcttatagcattattagcattcctttctatt
 ttgttcttaataaccgaaaaaacagtctaattgtataa

Sequence 2566
 15 VPLQSSYNSDIGTINIAVLSALFSGLFIVGAGDIADKIGRVKVTYIGLALNIVGSILII
 ITPLPSLLIIGRAIQGLSAACIMPATLAIINEYYIGTARQRALSYWSIGSWGSGVCTLF
 GGLMATNLGWRISFIVSIILTILSMFLIKHTPETKAEPIDQPTETKKFDVVGLIILVVS
 MLSINVIITQTSQFGLFSPFILGLIAIFVISLIIFVIYENKIKQPLVDFDIFKNKGYTGA
 TISNFMNLGVAGGTLIVVNTFYQQKLDNFSQETGYISLTYLIAVLIMIRVGEMILQSLGP
 20 KRPLLLGSALTIVIGLILLSLTFLPNAWYIASSVIGYLLFGTGLGVYATPSTDTAVAQAPD
 DKVGVASGVYKMASSLGNAFGVAISSTVYSVLAAQLNLTLGGFTGVMFNALIALLAFLSI
 LFLIPKKQSNV*

Sequence 2567
 Contig_0774_pos_4133_5074,
 25 putative peptide of unknown function
 atggaatatatgaaaaatagcaattgcaggttctggcgcattaggttagtgatttggtgct
 aagtgtttcaacatggttatgacgtcactttaattgataattgggaacctcaagtact
 acaatacaacaggacggtctacatatcgatattaatgggtgaagcgcatcatttcaggcta
 cctatgtatagactaacggaaattcctaagcaacgtcctatgatattgtttttctattt
 30 cctaaatctatgcaatagaagaggtgcttagtcatattcaaccccatcttcatagataat
 acaattgttgtgtgcactatgaatggtttgaaacatgaacgtcttatacaacaatatgtt
 tctatagatagaattgtacgtggagtaacaacgtggactgccggtattgatcaacctggt
 cacacgcacttaattggggcaaggtcctgttgaaattgggtgtctcgtgcccgaggga
 gaaacgctagatatcattgttaactctgctacaaaatgcagaattaaaagggtgtaaaaagt
 35 gaacatttacatcaatcaatttgggaagaaaatatgtgttaattggaacagctaattcatta
 tgtactatacttgaatgtaatttggcagcactgaataatagtgatgacgctaaaaatttg
 atatataaaattacacaagaaattgttcatgttgcaacagttgatgatgttcatctaat
 gttgatgagatttttgattacttaattgctttaaatgataaagtagggccacactatcct
 tctatgtaccaagacttaattaaagataatcgaacaactgaaatagattatattaatgga
 40 gcagttagtaaatagggaagagaatcatattgctacacctgtaaatgattttgtaaca
 aatcttgtagatgctaaagaaaatcaacgtggtgcacaatga

Sequence 2568
 45 MEYMKIAIAGSGALGSGFGAKLFQHGVDVTLIDNWE PQVTTIQDGLHIDINGEAHHFRL
 PMYRLTEIPKATSYDIVFLFPKSMQLEEVLSHIQPHLHDNTIVVCTMNLKHERLIQQYV
 SIDRIVRGVTTWTAGIDQPGHThLMGQGPVEIGCLVPEGKESVDIIVNLLQNAELKGVKS
 EHLHQSIWKKICVNGTANSLCTILECNLAALNNSDDAKNLIYKITQEIHVHATVDDVHLN
 VDEIFDYLIALLNDKVGPHYPSMYQDLIKDNRTTEIDYINGAVSKLGKENHIATPVNDFVT
 50 NLVHAKENQRGAQ*

Sequence 2569
 Contig_0774_pos_2605_2297,
 putative peptide of unknown function
 55 atgattaaaatcgaccaacaatatgaagtctaagcctatgtacgtcacttttactcta
 ccaattttatctgcaatatctcctgcacctaacaataaacagaccagaaaatagtgcgcat
 aagctaactgctatattaatagtaccaatatcactattgtatgatgattgtaattggcacg
 actagatttacaagtattgtgcaaataaccaaataatgtaataacccctaaaataatccct
 aataataagcgattatctcccctaaactttgtgatgtattcatctatcacttactcctt
 cacaattaa

Sequence 2570

MIKIDPTILSAKPMYVTFITLPILSAISPAPTINRPENSADKLTAILIVPISLLYDDCNGT
TRFTSDCANNQNVITPKIIPNNKRLSPLNFCDFIYHLLLHN*

5

Sequence 2571

Contig_0774_pos_0_597,

is similar to (with p-value 3.0e-65)

>gp:gp|U81973|SAU81973_16 Staphylococcus aureus capsule gene
cluster Cap5A through Cap5P genes, complete cds. NID: g1773
339.

atgaaaaaagtattgaccatatttggaaactaggcctgaagctataaaaaatggctccgttg
attaaaaacgttagagaaaagattctgacctggaacccggtgttgtagtcaccgcccaacat
agagagatgcttgattcagtggtgaatacttttaacataagtgacagattatgatttgaat
15 attatgaaagctgggtcaaacattgtctgaagtaacatctgaagcaatgaaaaagttagaa
gatatcatacaaaaggaagtgccatgatattggtacttgttcattggtgatacagtgacaacc
ttttctggagcatttagccgcattttatagtcacacacatagggacatggtgaagcggga
ttaaggagttataataaatattcaccttatcctgaagaaataatagacaaatggttggg
gtaatggcagatttgcactttgccccaacctataatgctgcacagaatttagtaaaagag
20 ggtaaatagccaaacatatagctatcactggtaatacagctattgacgcaatgaattat
acaatcgatcaccaatattcatcatctatcatacaaaaacataaaaaataaaaaTACT

Sequence 2572

MKKVMTIFGTRPEAIKMAPLIKLEKDSLEPVVVTAQHREMLDSVLNFTNISADYDLN
25 IMKAGQTLSEVTSEAMKKLEDIIQKEVPDMVLVHGDTVTTFSGALAAFYSSQTPIGHVEAG
LRSYNKYSPYPEEINRQMVGMADLHFAPTYNAAQNLVKEGKLAKHIAITGNTAIDAMNY
TIDHQYSSSIQKHKNKNT

Sequence 2573

30 Contig_0775_pos_5156_0,

putative peptide of unknown function

gtgtttatagctgctgaacatcaaacgtcagttgatgagttagagtatatattatgttcca
cctttaactacagaagctcaaaatgcctttgacgaacttaagcatgtgatgccaaaagta
tgttataaaagccaatttagaagcgtttggcatcggttgccaaatataaataacactgacat
35 aatcctaattgctgaagcgcacgtcacgctagcgattggagtgaaagttcgccagaatgg
ggtctagcacgaaatgctgaattcattatttgggaaacgtcaaatcacccaaaatagtaat
ctagagggacgggcattttcttcataattatgattggacaaaggatgaagacggtgagatt
ttaaatacaattattttctgggcagcactagtagcacaatggattaatttacaattactac
gcctcaaccgtggcacctcactattatggaagcggtagtaaaacaacgcaaacgtgaaca
40 agtgggtgtaggtgtcatgcaaggaaatgctagtatttaattgtatggcttaccatggcag
tcagtaattgatgaatgacaaaaggcggtatcacgcacctattaggcttttaattgttatt
caagcgccagatgcataatattcaacgtttgttaaaacatcataatcacttttagacaaaag
gttgatcatcaatggataagacttgccagattgatgaaaataatagttggaagactgg
ta

45

Sequence 2574

VFIAAEHQTSVDELEYIYVPPLTTEAQNAFDELKHVMPKVCYKANLERLASLPNINNTDH
NPNAEAHRHASDWSEVRPEWGLARNAEFIIGKRQITQNSNLEGRAFLHNYDWTKDEDGEI
LNTIISGPALVAQWINLQYYASTVAPHYYGSGSKTTQTVTSGVGVMQGNASDLMYGLPWQ
50 SVMMDKEAYHAPIRLLIVIQAPDAYIQRLKHHNHFRQKVDHQWIRLASIDENNSWKDW
X

Sequence 2575

55 Contig_0775_pos_3674_2475,

is similar to (with p-value 2.0e-92)

>sp:sp|P44953|DFP_HAEIN DNA/PANTOTHENATE METABOLISM FLAVOPRO
TEIN HOMOLOG. >pir:pir|G64104|G64104 pantothenate metabolism
flavoprotein (dfp) homolog - Haemophilus influenzae (strain
Rd KW20) >gp:gp|U32776|U32776_9 Haemophilus influenzae Rd s

ection 91 of 163 of the complete genome. NID: g1573969.

atgaacatattttatttagctgttacagcggtatcgagcatataaagcaattgattta
 acaagtaaattaatacaatccggctatgatgtaagagttatgctatctgatcatgctcaa
 5 gagtttggtactccgctagcttttcaagcaatcagtagaaatcctgtttacacaaataca
 tttaaagaagaaaatcctgaagagattcaacatgtatcattaggagactggcagatgcg
 attatagtcgcgccagcaactgctaatactatcgcaaaattaagtgttggaattgctgat
 gatttaattacttctacattacttgcatacaaacacacaaaattcgttgcacccgcaatg
 aatgtaaatatgtataacaatccacgtactaaacataatgaaagtgttaagtcaagac
 ggatattattttattgaacctggtagtggtatttagcatgtgggttatgtagcaaaagg
 10 cgaatggaagaacccatgcaaatcctatctgttattaataaaatttttactcaacagaag
 aatggtgtcaaaagctcttttctggaagcgcgcattagttacagctggcgctacagtt
 gaagttattgatcctgttcgatcgtatcaaatcggtcatcaggaaaaatgggatatgct
 atagctgaagcattacagagataaggagcaaatcgtaactttaattagtgggtccaccac
 ttatctctacctgaagggattaatgtagtaaaagttgagagtgcagatgatatgtttcaa
 15 gctgtaaccgaacgctttgcgaacaagatatagtgattaaagcagcgcggtgtctgat
 tataaccaatggacatacttgaacataaattaaaaaacaagaaggaggattatctgtt
 caatttaagcgtacaaaggatattttaaaatacttaggagaaaaataaaacgcaccaat
 ctgtgtggttttgcgtgctgaaacacaaaatattgaacagtatgctctagacaaactcaa
 agaaaaaatgcagatggttatcatttcgaacaatgtaggtgatacatccataggctttagt
 20 tcagatgacataaactaatatgcattttaaaaataatgaaaaagtaaatattaagaaa
 ggaaaaaatcagcttttagcacatcaaattatagaaattttagaaactaggtggcagtaa

Sequence 2576

25 MKHILLAVTGGIAAYKAIDLTSKLIQSGYDVRVMSLDHAQEFVTPLAFQAI SRNPVYNT
 FKEENPEEI QHVS LGDWADAI I VAPATANTI AKLSVGIADDLITSTLLATTPKFVAPAM
 NVNMYNNPRTKHNMKVLSQDGYFYIEPGSYLACGYVAKGRME EPMQILSVINKFFTQOK
 NVVKSSFSGKRALVTAGPTVEVIDPVRYVSNRSSGKMGYAIAEALRDKGAI VTLISGPTH
 LSI:PEGJ:NVVKVESADDMFQAVTERFAKQDIVIKAAVSDYTPMDILEHKLKKQEG:GLSV
 30 QFKRTKDILKYLGENKTHQYLVGFAAETQNI EQYALDKLKRKNADVIISNNVGDTSIGFS
 SDDNELTMHFKNNEKVNIIKKGKKSALAHQIIEILETRWQ*

Sequence 2577

Contig_0775_pos_2367_67,
 35 is similar to (with p-value 0.0e+00)
 >sp:sp|P94461|PRIA_BACSU PRIMOSOMAL PROTEIN N' (REPLICATION
 FACTOR Y) (FRAGMENT). >gp:gp|Y10304|BSPRIADFS_1 B.subtilis p
 riA, def, fmt, sun genes. NID: g1772497.

gtgggtacctttggccctagaactatacaaggatacgtcatgaatattcaacaaaaacca
 40 gtggaaatatggatatatcgaaactaaaagaaataaaagaagtacgtgatattaacct
 gaattaacatccgaactgattcaattaagcgaatggatgagccattatcatgtgatgaaa
 cgtattttctgttttagaagcgatgttgccaagtgccattaaagcaaagtataagaaagct
 ttttcaattatcgatccaaaaaatttatcttcaaaaaccaaagcgctatttaacaatgac
 ggttattacttatataaagaagttcagcaaaacaatgatttagaagaaatgttgactttg
 45 ttaaatcaaggattgattgaagaggtcacgatactttctcaaaacacaaaaagaaaact
 caaaaagctgttgcgtagttaatacgttggaatgggtgatgaagtacttgcaaaactcgag
 aaatatacaaaaacaatatgatttgtatgcatttttattagaagagtcctcatcgaaagtg
 tttttaaagaagaatcaatgatatgggcttctctcactcgagtttagattctttaataaaa
 aaaggtctattgaaaaatatatcgccgaggttttcagagatccatagcaaatcgtata
 50 tttgaacaagaacaaaagaggatattaactaaagaacagcaagatgcatttgaagctatt
 caacattatattcatgatgaaaaagaagaacatttttattacacggagtcacagggttca
 ggtaaaaccgaagtctatcttcaacaatagaagaagttcttaataaaaggtaaagaagcc
 atgattgttagtgcggaaatcgcttaacacctcagatgggtactaagatttaaacgtcga
 tttggagatgatgtagcgggtattacattccggactttcaaaaggtgaacgttatgatgag
 55 tggcaaaaaattagagacggtcgagctcgagtgagtgtaggtgctcgttcaagtattttc
 gcaccgtttaaaaatttaggcattatataattgatgaagaacatgaatctacatataaa
 caagaagattatcccagatatcatgcacgtgatattgcacaatggagaagtcaatttcat
 cattgtcctgtagtttttaggtagcgcaacaccgagcttgagtcatatgcaagagcagaa
 aaaaatgtttacgagttgttgtcattgccacatagagtcaatcaacaagcgttaccgcat

atcgatat.tatagatatgagagaagaattaagtgaaggtaatcggtccatgtttttatata
 gcactaagacaagcgatacaagaacggttgataaaaaagaacaaatagtactattctta
 aatagaaggggatatgcttcatttatgttatgtagagattgtgggtacgttcccaatgt
 ccccatgtgatatttcgttaacatatcataaaacaaccgatcaattaaaatgtcattac
 5 tgtgggttatcaagaaaatccaccatctcaatgtccaaattgtgaaggatcatatcaga
 caagtcggaactggaacgcaacgtgtagaagaattattacaacaagaattccctcatgct
 cgtattataaggtatggtatgatacaacttcaagaaaagggtgcacatgagaaattggtta
 aatgactttgaagcaggaagagatatcttattaggtacgcaaatgattgctaaagggt
 ttggattatcctaacattactctagttgggtgtgctcaatgctgacactatgttaaactta
 10 cctgactttcgtgccagtgaacgaacataccaacttttaactcaggtatctggacgcgca
 ggtcgtcatgaaaaagaaggacaagttatcatacaaacgtacaaccctgatcattattca
 ataaaggatgtgaaattaaacgattatctttctttttatcaaaaaagaaatgaattatcga
 aaattaggaaaaatccaccttactattttttgattaaactttaccatttcacatactgat
 ataaaaagggtcatggtgctcctaagcatatacatcaaatttttagtacagcacttaagt
 15 gaaaaagcattcgtgcttaggcccttcaccagcagcactagcaagaattaaacatgagtat
 cgttttcaaatactagtataaatataagagtgaagcctcaattacatcaagcgttacaatat
 ttagatgattataatcatgatcaatatgtaaaggataaactatcattaaaaattgatata
 aatccacaaatgatgatgtga

20 Sequence 2578

VVPEFGPTIQGYVMNIQOKPDGNMDISKLEIKEVRDIKPELTSELIQLSEWMSHYHVMK
 RISVLEAMLPSAIKAKYKKAFSIIDPKNLSSKTKALFNNDGYLYKEVQQNNDLEEMTL
 LNQGLIEEVTILSQNTKKKTQKAVGVVNTLNGDEVLAKEKYTKQYDLYAFLEESHRTV
 FLKEINDMGFSHSLDSLKKGYIEKYIAEVFRDPYANRIFEQEQRILTKEQQDAFEAI
 25 QHYIHDEKERTFLLHGVGTSGKTEVYLQTIIEVLNKGKEAMMLVPEIALTPQMVLRFKRR
 FGDDVAVLHSGLSKGERYDEWQKIRDGRARVSVGARSSIFAPFKNLGIIIDEHESTYK
 QEDYPRYHARDIAQWRSQFHHCVPVLGSATPSLESYARAENKVVYELLSLPHRVNQALPH
 IDIIDMREELSEGNSMFISIALRQAIQERLDKKEQIVLFLNRRGYASFMLCRDCGYVPQC
 PHCDISLTYHKTTDQLKCHYCGYQENPPSQPCNCEGDHIRQVGTGTQRVEELLQQEFPHA
 30 RIIRMDVDTTSRKGAHEKLLNDFEAGKGDILLGTQMIAGLDYPNITLVGVNLADTMLNL
 PFRASERTYQLLTQVSGRAGRHEKEQVLIQTYPDPHYSIKDVKLNLYLSFYQKEMNYR
 KLGKYPYFLINFTISHTDIKKVMMASKHIHQILVQHLSEKAFVLGSPSPALARINNEY
 RFQILVKYKSEPQLHQALQYLLDDYNHDQYVKDKLSLKIDINPQMM*

35 Sequence 2579

Contig_0776_pos_1350_2054,
 is similar to (with p-value 2.0e-50)
 >sp:sp|P54163|YPDP_BACSU HYPOTHETICAL 25.7 KD PROTEIN IN BCS
 A-DEGR INTERGENIC REGION. >gp:gp|L77246|BACYACA_10 Bacillus
 40 subtilis (YAC10-9 clone) DNA region between the serA and kdg
 loci. NID: g1256615. >gp:gp|Z99115|BSUB0012_140 Bacillus su
 btilis complete genome (section 12 of 21): from 2195541 to 2
 409220. NID: g2634478.
 atgtataatgaaatatttggtattgcgtcatttattgttacattcgctttaatggtagtgc
 45 atgtatcgctgttttggtaaacaaggactaattgcttgggttagcaataggaacgattatc
 gctaataacaggtcataaaagcgggtcatatttttggtattacggctacacttgaaat
 gtcattgttgcctctatatatttagctactgatataataatgacatctatggctgtaaa
 gttgctaaaagagcgggtgtggttctctcttctaccttagtaattgattatagtcata
 caaatgtcattgcattttattcctgccccagtagacaatgcgcaaaactcattaaaaatg
 50 atttttgatttagtgcttagaattgctatagggttcattattgcttatatcataggccaa
 catattgatgtatttatattcagtatgattaaaaagatatttagctctgataagaccttt
 ttatttagagcatatggtagtaccattttaagttctatcattgataccggtttatttggt
 tcaattgcttttattggtactatgcctggtactgctgttttgaaatatttattaccact
 tacttgtaaactagtgtcaactatttttaattgaccatttgatatatcgctaagtca
 55 ctatatcgaaaaggaaagatagaacaactagataatgggtattga

Sequence 2580

MYNEIFGIASFVTFALMVLMYRCFGKQGLIAWVAIGTIIANIQVIKAVHIFGITATLGN
 VMFASIYLATDILNDIYGRKVAKRAVWLGFSSTLVMIIVMQMSLHFIPAPVDNAQNGLKM

IFDLVPRIAIGSIIAYIIGQHIDVFIFSMIKKIFSSDKTFFIRAYGSTILSSIIDTGLFV
SIAFIGTMPGTAVFEIFITTYLLKLVTIFNVFPGYIAKSLYRKKGKIEQLDNGY*

Sequence 2581

5 Contig_0776_pos_2379_2771,
is similar to (with p-value 3.0e-19)
>sp:sp|P36921|EBSB_ENTFA CELL WALL ENZYME EBSB. >pir:pir|B49
939|B49939 ebsB protein - Enterococcus faecalis >gp:gp|L2380
2|ENEEBSA_2 Enterococcus faecalis pore forming, cell wall en
10 zyme, regulatory, and dehydroquinase homologue proteins (ebs
A, ebsB, ebsC, and ebsD) genes, complete cds with repeat region
. NID: q388106.
atggctaaaatacattttgatgctgcgactaaaggaaatcccgccgaagtgcttgtgcg
attattattaaagaaaattcacaaagatatatcattttacccatgatttagtgaaatggat
15 aatcatagtgacagaatggcgagcaatgttacacgcttttgaacatgcacgcgaattaaaa
gtatctaacgcgcttactttttactgattcaaaattaattgaagatagtatgatgcaaggt
aaagttaaaaatgctaagtttaaagtttattttgaaaacatagaaatcttagagcaaagt
tttgatttgatgtttgtgagatggattccacgaaagcaaaataaagaagcgaatcaact
gctcaacaaacactatacaaaacttacatcataa

Sequence 2582

MAKIHFDAATKGNPGRSACAI I I KENSQRYTFT HDLGEMDNHSAEWAAMLHALEHARELK
VSNALLFTDSKLIEDSM MQKVKNKFKVYFENIEILEQSFDLMFVRWI PRKQNKANQL
AQQTLYKLTS*

Sequence 2583

Contig_0776_pos_5545_4955,
is similar to (with p-value 3.0e-22)
>gp:gp|AL034447|SC7A1_23 Streptomyces coelicolor cosmid 7A1.
NID: g4007715.
gtgttcggttatggaaaaagaagaatctttcttctgaaagtgaaaaaatgggaacct
atcgacttattttttatatggtttgcagccaatctaggtattctcggtatagtttatggc
gctgtcatagtaagttacggattgagctttttacaatcaattttgattgctataatagga
ccgttatccctttgttctagttggttatataagtgtagctggtagagatagtgagctata
35 acttttatgctctcaagagcaccatttggatttaaaggcaatcatataacctgctttaatt
ggctgggtaggtcaagttggttggttatctgttaattgtttctacaggaactttaactctt
ctggctttattcaataacttttggttttaagactagtagctatttctaattttgatgagttta
gcgatttttgcgtgggctagttattatatctgttctttttcaaaaagtagcttgatca
gtacaaacatttttcacatatgtatttgggtgcattaaccttattagttataacaatttta
40 attactaatactgattggaacgcccttttttctatgaaatctgggtcttaa

Sequence 2584

VFVMEKEEIFLPESEKNGKPIDLFFIWFANL GILGIVYGAVIVSYGLSFLQSILIAIIG
PLSFVLVGYISVAGRDSGAITFMLS RAPFGFKGNH IPALIGWVGQVGWLSVNVSTGTLTL
45 LALNFTFGFKTSTFLILMSLAIFAGLVII SVLFSQKVLVSVQTFFTYVFGALTLLVITIL
ITNTDWNALFSMKS GS*

Sequence 2585

Contig_0776_pos_4784_2949,
50 putative peptide of unknown function
gtgataaaagagttatcacaaaagaaacgagatgcaataaataacaacactgatttaaca
cctttctcaaaaggcacatgcttttagcagatattgataaaacagaaaaagatgcac tcaa
catatcgaaaattctaattcaattgatgatatacaataacaataaagagcatgcatttaatt
actttagctcatatcattatttgggatactgatcagcaaccattagtttttgaaactacct
55 gaattgaccttcaaaatgctctagtaacaagtgaggtggttgttcacagagatgaaact
atttcattagaatctataattggagctatgactttaactgatgaacttaaagtcaatatt
gtttcattaccgaacactgataaagtagctgatcacctaaccgctaaagttaaggttatt
ttagctgatggctcatttgtcactgtaaatgttccagtc aaagttgtagaaaaagaatta
caaatagctaaaaaggtgctataaaaacaattgatgttctggtaaaaacaaaaaatcaaa

gatatagattctaataacgaattaacgtctactcaacgtgaagatgcaaaagctgaaatt
 gaaagattgaaaaagcaagccatcgataaagtgaactcattctaatacgattaaagatatt
 gaaacagtaaaacgaactgattttgaagaaatagatcagtttgatcctaaacgctttacg
 ctaaaataaagctaaaaagcatatcattactgatgttaatactcaaatccaaatgctttc
 5 aaagaaattgaaacaataaaagggtttaacttctaatagaaaaactcagtttgataaaca
 ttaactgcactacaaaaagaatttttagaaaaagtcgagcatgctcataattagtagaa
 ttaaatcaattacaacaagagtttaataatagatatgaacatattttaaccaagcacat
 ttactaggtgaaaaacatatagcagaacataaattaggtatgttgtagtaaacaaaact
 cagcaataactaaataatcaatctgcttcttactttataaaaacaatgggcacttgataga
 10 attaaacaaattcaactagaaacgatgaattcaattcgtggcgcataccgtacaagat
 gtacacaaaagcattgttacaaggatatagagcaaatcttgaaagtaaatgtaagtattata
 aatcaatctttcaacgattccttgcataactttaattatcttcattcaaaatttgatgct
 agattaagagaaaaggatgttgcaaacatatcgtaaaaactgaaacattcaaagaagtt
 ctaaaagggaacgggtgttgaaaccaggtaaaatcaacaaagaaacacagcaacaaaactt
 15 cataagaatgataatgatagcctattcaaacatttagttgataatttcggcaaaaactgta
 ggtgttattacattaactggttactttctagtttctgggttagttttggctaaaagacgt
 aaaaaagaagaagaaaaacaatcgataaaaaattatcacaagacattcgtctttcagat
 actgataaaaatagatccaattgttaataactaagcgtaaaatagataaagaagaacaaatt
 caaacgatgacaaacattcaattccagttgctaaacataagaaatctaagaaaaagcaa
 20 ttgagtgaaggagatattcattcaatccccgtcgtaaacgtaaacaaaacagtgataac
 aaagatacaaaacagaagaaagttacttctaaaaagaagaaaacgcctcaatcaactaaa
 aaagttgtaaaaacaaaaagcgttctaaaaagtaa

Sequence 2586

25 VIKELSQKKRDAINNNTDLTPSQKAHALADIDKTEKDALQHIENSNSIDDINNKEHAFN
 TLAHIIIWDTDOQPLVFELPELSLQNALVTSEVVVHRDETISLESIIGAMTLTDELKVNI
 VSLPNTDKVADHLTAKVKVILADGSFVTNVNPKVVEKELQIAKKDAIKTIDVLVKQKIK
 DIDSNNELTSTQREDAKAEIERLKKQAIDKVTHSKSIKDIETVKRTDFEEIDQFDPKRFT
 LNKAKKDIITDVNTQIQNGFKEIETIKGLTSNEKTQFDKQLTALQKEFLEKVEHAHNLE
 30 LNQLQQEFNNRYEHILNQAHLLGEKHIAEHKLGYYVVVNKTQQILNNQSASYFIKQWALDR
 IKQIQLETMNSIRGAHTVQDVHKALLQGIEQILKVNVSIIINQSFNDSLHNFNYLHSHKFD
 RLREKDVANHIVQTETTFKEVLKGTGVEPGKINKETQPKLHKNDNDSLFKHLVDNFGKTV
 GVITLTGLLSSFWLVLAARRKKEEEKQSIKYNHKDIRLSDTKIDPIVITKRKIDKEEQI
 QNDDKHSIPVAKHKKSKEKQLSEGDHISIPVVKRKQNSDNKDTKQKKVTSKKKKTPQSTK
 35 KVVKTKKRSK*

Sequence 2587

Contig_0776_pos_587_147,

is similar to (with p-value 9.0e-29)

40 >sp:sp|P54170|YHPH_BACSU HYPOTHETICAL 15.9 KD PROTEIN IN ILV
 D-THYB INTERGENIC REGION. >gp:gp|L77246|BACYACA_21 Bacillus
 subtilis (YAC10-9 clone) DNA region between the serA and kdg
 loci. NID: g1256615. >gp:gp|Z99115|BSUB0012_128 Bacillus su
 btilis complete genome (section 12 of 21): from 2195541 to 2
 4092220. NID: g2634478.
 45 atgatgaatggatacgaagcttatatgaaagaacttgacacacaaatgagagctgaatta
 acagacaatggattcacaagcttgaaacgagtgatgacgtcaatcagtatatgcaaaat
 atagataatgatgatacaacatttgggtgaatcaactcaacatgcggttgcgctgcagga
 ttgacacgtccagcagctgttgagttgcagagcaaaatgaagtgaaccagatcataaa
 50 gtaactgtatttgcgtggtcaagataaagaagcaacacaaacaatgagagattacatccaa
 caagttccttcaagtccttcatacgcattatttaaaggtcaacatttagttcattttata
 cctcgcgacatattgaaggcgcgacatcaatgatatagctatggatttaaaagatgct
 tttgatgataattgtcaataa

Sequence 2588

55 MMNGYEAYMKELAQQMRAELTDNGFTSLETSDDVNQYMQNIDNDDTTFFVINSTCGCAAG
 LARPAAVAVAEQNEVKPDHKVTVFAGQDKEATQTMRDYIQQVPSSPSYALFKGQHLVHFI
 PREHIEGRDINDIAMDLKDAFDNQC*

Sequence 2589

Contig_0777_pos_367_684,

putative peptide of unknown function

5 atgaggagaatatatatgaaaagattattaggtacattaattgctgctacactagtgtta
 agtgctttagccaaaacgacactaaggaagatgaaaataaaaagtcagaaaatactact
 gaaaagaatctgacgataaaaaagacaaaaaactaatgaggataaaaagtcaggagaa
 caaaagaatctcaagaaaaaagaataacaagtcaatgcaagaatctgctacaaatgaa
 caggttcaatctcaacaacaacttcacaagggtgtaccatcaaaatgctctttgccgatg
 actgttttatcaatttga

10

Sequence 2590

MRRIYMKRLLGLTIAATLVLSACSQNDTKEDENKKSSENTTEKKSDDKKDKKTNEDKKS
 QKKSQEKKNKSMQESATNEQVQSQQQTSQGVPSKCSLPMTVLSI*

15 Sequence 2591

Contig_0777_pos_6354_6046,

putative peptide of unknown function

atgactgcatttttaattcttaattttattgactatgaaagaggcttcaatatcgtctatt
 attgtacgaactgtcattgcagctattgtttctttgtcatttatatcattgtatttaca
 20 attttaagttcgtcagaacgtaaaccttatttatgggtacaactttgcctattgcgctttt
 atatgccttatattcggagcaatttttttctactccgcgtataggtatcattgccggacta
 attataggtgtgtttgctggtgtcatatgggagttcttaaatagaaaaatggaggtcgc
 tcatcttga

25 Sequence 2592

MTAFLILNLLTMKEASISSIIIVRTVIAAIVFFVIYIIIVFTILSSSERKLIYGTLPIALF
 ICLIFGAIFFTPRIGIAGLIIGVFAGVIWEFLNRKNGGRSS*

Sequence 2593

30 Contig_0777_pos_4678_1709,

is similar to (with p-value 0.0e+00)

>sp:sp|P40815|T3RE_SALTY TYPE III RESTRICTION-MODIFICATION S
 YSTEM STYLT1 ENZYME RES (EC 3.1.21.5). >pir:pir|JN0658|JN065
 8 restriction endonuclease (EC 3.1.-.-) - Salmonella typhimu
 35 rium

atgaaaaatattcttagaagaactaattcaccaacaacaagcagttaaaaaataatagat
 actttcacaggaatcgaaaagtatttaacttcaaaaaattgtgacaatgagttcgcta
 aatttaataataatagatatattcagaagaagctaacatagatataaaaatggagaccggg
 acaggtaaaacgtatgtttataactaaaatgatgtatgaattacataaaaaatttgggatt
 40 ttttaatttatatttagtagttccaagtccctgcaattaaagagggagcaagaatttttta
 actagtttatcaactaaaagacatttccaagaaacatacggaaatgttgaaatagaaata
 aatacaataaataaaggcgatttttaatactagatcaggcagaaaaattttccacctcat
 ttgagtaactttattgaaagcagtaatttaaatgctaatacaattcaagtattacttatt
 aatgcagggtatgttaattcatcaaatatgacaaagggtgattacgaccaaactttatta
 45 agtaattactcaaaccttatagaagctttaaagctaccaaactctgtagtataatagac
 gaaccccataggtttcctagggataagaagaactacaaatctatagaaaatcttgaacct
 caaatgattgttaggtttggtgctacttttcctgaagtaaaaaaggaacaggaaaaaaa
 gctgtatatattaaagattattatcgaggtaggcctcaatttgaattgaatgctgttgac
 agctttaatcagggaactgttaaagggaatagatatattactatccaaatttgactccagaa
 50 caagcgaaaaatagatataccattgatagtgtaaaagccaaagagatagttttgaaaaaa
 ggaaaaaataaatggagcgttaggtataggcgagaatttagcaaatatagattcggttattt
 gaaggtgacttaagtattctggtgctaaaactttatctaattgacttagaaattagtaaa
 ggtatgattttgtcctggcacatttacaactaattaccaggaatttaattattaatgat
 gctatcaatcaacattttgagcatgaaatcaataattttatgagagataacctaanaagaa
 55 aattttcaaccaaagtaaaaacactatcattattttttattgattctattaggtatata
 agaacaaagagggtggttgaaacagacttttgagagactcttaaaagttaaacttaga
 aaattaataaaaagaatttgaagttaaaaaattaccacgagagattgaatacttagatttt
 ttaagagtgcacatatgatagtccttaatagtgagaatcaaatggccatgcagggtatttt
 ggtgaagacagaggttccgggtgatgaagcaatacagggtgaagttaatgatataacttaa

aataaagaaaaatgctgagttttaagacgagaatggtaactggattacaagaagattt
 ttattttcaaaatggacattaaagagaaggttgggataaccctaattgttttcgttattgcg
 aaattgagaacttctggttctgaaaatagcaagatacaggaagtaggaagaggtttgcgt
 ttacctgttgatgagaatggtcacagattaacacaagatgaattccaagtagactttca
 5 tttttaattggatatgatgaaaaggatttcgcagaaaaattaattggagaaataaattct
 gatgtagacgttaaattaagtgaagataaattaacagatgaaatgattaataaaatagtt
 gaacatagaaagcaagtggatcctcattataatgatgaagtgttattagaacaattagat
 gaacgaaatctaattaataggaaaaatgagtttaaacagatgttgaaattgatggagtt
 aaaaaatcaggatttgaatggctgttagaattatatccagaagttaatgcatctaaatta
 10 aatgcagataaagtgcagagatatgaaaaaaaatccaccaatttaaaagtgaagttaaat
 aaagagaattggaataaattgagattccttatgggaaaaatttatctaaaagatatatgta
 gaattcaaaaagatgagtgaaatgatctgtatttgtttgtgaaagattgttgaacgac
 gatgattttattgttaacaacaaccagaaagaatacatcaatctttggaaaaagatgat
 gaaggtaaacaagttatcaaggaatctatttcagaatataattatagaaatgaatttatt
 15 tatatgaactatggtaagttttaaaacaaattgtacttaaaactaatgtaccaatatct
 ataatgcataaaaatttactttccggttttaaaagataaatataatagtgtgaacgattc
 ttgagtgaattaagtttaataatataataagagaatttaacaagcgttttgaagaaaaa
 tatagctaaagttatgaatataagaaattagatttttctgctactacaaccatttatgat
 tcagaaatcatcagagtttaagattgggtagatgcaaattatttaggtactaacgttgaa
 20 aataacattcaaaactgaaaaaagatttttatatgaaagaccaccagttagatatgatagt
 gtaacacctgagtttagagttgttaaaaagaaattacgataaaaatgtaactgtatttgggt
 aatttgcctaaaaaagcgatacaagttcctaataatactggtggcactactacgcctgat
 tttgtctatatgatagaactgatgaacaagatgcaaaataccttattgttgaaacaaaa
 gcagaaaacatgagactaggagataaaagtattggtgaaatacaaaaaaaattctttaac
 25 acattagataaatttgaattataaatatcaatttagctactagcgcgcaagatgtttataat
 gaaattaaaaaatttagatgattcaagtgatga

Sequence 2594

MKIFLEELIHQQQAVKKIIDTFTGIEKYLTSKNCDNEFANNLIINRYSEANIDIKMETG
 30 TGKTYVYTKMMYELHKKFGIFKFILVVPSPAIKEGAKNFLTSLSTKRHFQETYGNVEIEI
 NTINKGDFNTRSGRKIFPPHLSNFISSNLNANQIQVLLINAGMLNSSNMTKVDYDQTL
 SNYSNPIEALKATKSVVVIDEPHFRFPROKKNYKS IENLEPQMIVRFGATFPEVKKGTKKK
 AVYIKDYRGRPQFELNAVDSFNQGLVKGIDIYYPNLTPEQAKNRYTIDSVKAKEIVLKK
 GKNKWTLGIGENLANIDSLFEGDLSYSGAKTSLNDLEISKGMOLLPGTFTTNYQELIIND
 35 AINQHFHEINNFMRDNLKENFQPKVKTLSLFFIDSIRS YRNKEGWLKQTFERLLKVKLR
 KLIKEFEVKKLPREIEYLDFLRVTYDSLNSNQMVHAGYFGEDRGSGDEAIQAEVNDILK
 NKEKMLSFKDENGWITRRFLFSKWTLREGWDPNPNVFI AKLRTSGSENSKIQEVGRGLR
 LPVDENGHRLTQDEFPSRLSFLIGYDEKDFAEKLIGEINS DVDVKLSEDKLTDEMINKIV
 EHRKQVDPHYNDEVLLQLDERNLINRKNEFKTDVEIDGV KKS GFEWLLELYPEVNASKL
 40 NADKVRDMKKNPNNLKVKNKENWNKLRFLWENLSKRYM LEFKMSEDDLYLFVERLLND
 DDLFVKQQPERIHQSLEKDEGKQVIKESISEYNYRNEFI YMN YGKFLKQIVLKTNPVPI
 IMHKNLLSVLKDKYNSDERFLSELNLNIIREFNKRFE EKYSQSYEYKKLDFSATTTIYD
 SEISEFKDWVDANYLGTNVENNIQTEKRFLYERPPVRYDS VTPLELLKRN YDKNVTVFG
 NLPKKAIQVPKYTGTTTPDFVYMIETDEQDAKYLIVETKA ENMRLGDKSIGEIQKKFFN
 45 TLDNLNIKYQLATS AQDVYNEIKKLDDSK*

Sequence 2595

Contig_0777_pos_1393_623,
 is similar to (with p-value 0.0e+00)
 50 >gp:gp|U92974|LU92974_13 Lactococcus lactis unknown gene, p
 artial cds, and HisC (hisC), unknown, HisG (hisG), unknown,
 HisB (hisB), unknown, HisH (hish), HisA (hisA), HisF (hisF),
 HisIE (hisIE), unknown, unknown, LeuA (leuA), LeuB (leuB),
 LeuC (leuC), LeuD (leuD), unknown, IlvD (ilvD), IlvB (ilvB),
 55 IlvN, IlvC (ilvC), IlvA (ilvA), AldB (aldB) and aldR (aldR)
 genes, complete cds. NID: g2565137.
 atgggattatagagtactacttttattataaatatgtaactatagatgacctgaaactttt
 gcagccgaacattttgaaattttgtaaggaacatcattttaaaggaagaataactagtttca
 acggaaggcattaatggaacattatctggaacaaaagaagatactgataaatatatagag

catatgcatgcagatagtcggttttgctgatttaactttttaaattgatgaagctgaaagt
 catgcgttttaaaaagatgcacgtgcgtccaagacgtgaaattggttgacttgacttagaa
 gaagatattaatccacgtgaaattaccggtaaatactattctcctaaagaatttaaagcc
 gcactagaagatgaaaatactggtatatttagatgctcgaaatgattatgaatacagattta
 5 ggacatttcctggtgagctatttcgtcctgatataacacgattccgtgacttacctgaatgg
 gtgcgtaataataaagaacaactcgacggaaaaaatattgtcacatattgtacaggtggc
 attcgttggtgaaaaattttctggttggttagtaaaagaaggatttgaaaacgtaggtcag
 ttgcatggtggtattgtctacatacggtaaaagaccctgaaactaaagggctatattgggat
 ggtaagatgtatgtatttgatgaacgtatttagtgtcgatgtgaatcaaattgataaaaca
 10 gtcacggaagagcattttgatggtacaccttgtaagttgtgtgtga

Sequence 2596

MDYRVLLYYKYVTIDDPETFAAEHLKFCKEHLKGRILVSTEGINGTSLSGTKEDTDKYIE
 HMHADSRFADLTFKIDEAESHAFFKMHVRPRREIVALDLEEDINPREITGKYSPKEFKA
 15 ALEDENTVILDARNDYEYDLGHFRGAIRPDITRFRDLPEWVRNNKEQLDGKNIVTYCTGG
 IRCEKFSGWLKKEGFENVQLHGGIATYGKDPETKGLYWDGKMYVFDERISVDVNQIDKT
 VIGKEHFDGTPCEVCC*

Sequence 2597

20 Contig_0778_pos_324_815,
 is similar to (with p-value 5.0e-52)
 >sp:sp|P38424|YSXC_BACSU HYPOTHETICAL GTP-BINDING PROTEIN IN
 LONA-HEMA INTERGENIC REGION (ORFX). >pir:pir|S45102|S45102
 hypothetical protein X - Bacillus subtilis >gp:gp|X76424|BSL
 25 ONLA_3 B.subtilis lon gene for protease La. NID: g496556. >g
 p:gp|Z99118|BSUB0015_84 Bacillus subtilis complete genome (s
 ection 15 of 21): from 2795131 to 3013540. NID: g2635200. >g
 p:gp|Z75208|BSZ75208_85 B.subtilis genomic sequence 89009bp.
 NID: g1769994.
 30 atgaatataaatttttaataatattaacttaattataagtgtgtgtaaaaaaagcacagtat
 cctgacactggattaacagaagtagcggttaagtggacgctcaaattgtagggaaatctaca
 tttattaatagtagtattggcgtaaaaaatattggcgagaacgctcacaacaacctggtaag
 acacagacattgaatttctataatatagatgaacaacttattttgttgatgtaccagga
 tatggatacgcctaaagtaagtaagttcaacgagaaaaatttggtaaaatgattgaagaa
 35 tatattacacaacgagagaattttaaacttgttattcaacttgtcgatttaagacatcaa
 cctactgaagatgatgtgcttatgtacaattatcttaacattttgatataccaacactt
 gtaatatgtactaaagaagataaaaattgccaaggaaaagtagcagggtccagagaggct
 ccgaagatatga

Sequence 2598

40 MNINFNNINLIISAVKKAQYPDTGLTEVALSGRSNVGKSTFINSMIGRKNMARTSQPGK
 TQTLNFYNIDEQLIFVDVPGYGYAKVSKVQREKFGKMI EYITQRENKLKLVQLVDLRHQ
 PTEDDVLMYNYLKHFDIPTLVICTKEDKIAKGKVRGSREAPKI*

Sequence 2599

45 Contig_0778_pos_2037_3047,
 is similar to (with p-value 8.0e-80)
 >sp:sp|P41006|PYRP_BACCL URACIL PERMEASE (URACIL TRANSPORTER
). >pir:pir|S38893|S38893 uracil transport protein - Bacillu
 50 s caldolyticus >gp:gp|X76083|BCPYRQP_2 B.caldolyticus (DSM40
 5) pyrR, pyrP and pyrB (partial) genes. NID: g431229.
 atgctggttgcatattttatgagtggttaattgtacgtgattataggtattttcattaaa
 ttgagtggaacacattggttaattgcacttggtaccaccagtagttgtcggaccagtaata
 atggtcattgggttaagtttagctcctacagcagtaaacatggccatgttcgaaaattct
 55 gctgaaatgaaagggtataacttaagttacttaattgttgctttgattacattagcagta
 accatcatcgtccaaggattcttcaaaggatttttatcactaatacctgtacttataggt
 attatagtggtatattgtatccattttcatgggcatagttaaatttgctccaatagca
 caagcgaaatggatagattttcctcatattttatctaccatttaaagattacacaccatct
 tttcatttaggactcattctcgtgatgatacccggtggtgtgtgtgacggtgaagtgaacat

attgggtcatcaaatggtaattaataaaatagtaggacgcaatttctttgaaaatccaggt
 ttagataaatcaatcattgggtgatgggtgtttcaactatgtttgcaagtatgataggaggt
 cctcctagtagacaacttatgggtgaaaatataggtgtactagcgatcaccaaaatatatagt
 atttacgttattgggtgggtgaggcagttatagctatcattcttgcatttattggtaagttc
 5 actgctttaatatcttcaatacctacgccagtgatgggtgggtgtctcaattttattattc
 ggtattatagcagctagtggtttaagaatgcttggtgaaagtcaagtagatttcgcaagc
 aatcgcaacttggttatagcatcagttgtgcttggtgctgggattggtaatcttcttctc
 aattttaaaggcataggtatcaatttacaattgaaggaatggcattatcagcactttca
 ggaataatattaaatttaattttgcccagaataaaaaccaaataaattaa

10

Sequence 2600

MLVALEFMSGLMYVIGIFIKLSGTHWLMHLLPPVVVGPVIMVIGLSLAPTAVNMAMFENS
 AEMKGYNLSYLIVLITLAVTIIIVQGFFKGFSLIPVLIGIIVGYIVSIFMGIVKFAPIA
 QAKWIDFPHIYLPFKDYTPSFHLGLILVMIPVVFVTVSEHIGHQMVINKIVGRNFFENPG
 15 LDKSIIGDGVSTMFASMIIGGPPSTTYGENIGVLAITKIYSIYVIGGAIVIAIILAFIGKF
 TALISSIPTVMGGVSILFGIIAASGLRMLVESQVDFASNRNLVIASVVLVVGIGNLLI
 NLKGIGINLQIEGMALSALSGLIILNLILPKDKNQIN*

Sequence 2601

20 Contig_0778_pos_3073_3954,
 is similar to (with p-value 2.0e-90)
 >sp:sp|P05654|PYRB_BACSU ASPARTATE CARBAMOYLTRANSFERASE (EC
 2.1.3.2) (ASPARTATE TRANSCARBAMYLASE) (ATCASE). >pir:pir|A25
 015|OWBSAC aspartate carbamoyltransferase (EC 2.1.3.2) catal
 25 ytic chain - Bacillus subtilis >gp:gp|M13128|BACPYRB_1 B.sub
 tilis pyrB gene encoding aspartate transcarbamoylase, comple
 te cds. NID: g143383. >gp:gp|M59757|BACPYROP_3 Bacillus subt
 ilis pyrimidine biosynthetic (pyr) gene cluster (pyrR, pyrP,
 pyrB, pyrC, pyrAA, pyrAB, pyrD, pyrF and pyrE) genes, compl
 30 ete cds. NID: g387576. >gp:gp|Z99112|BSUB0009_20 Bacillus su
 btillis complete genome (section 9 of 21): from 1598421 to 18
 07200. NID: g2633902.

atggaacacttattatcaatggagcatttatctaattcagaaatttatgatttaattact
 atcgcttgccaattcaaatctggtagcgaccattacctaatttaacgggtcaatacgt
 35 tcaaaacttattcttcgaaaattcaacgcgaacaaagtgtagctttgagatggcagaacaa
 aaattaggattaaaacttattaattttgaaacaagtacatcatctgtaaaaaagggtgag
 tcactttatgacacatgtaaaacacttgaaagtataggtgttgatttacttgcatacgt
 cactcccaaaattcttattacgaagaactggatcaattaaatattccaattgctaatacga
 ggtgatggaagtggacaacatcctactcagagtttattagacataatgacaatatatgaa
 40 gaatatggttcgtttgaaagtttgaatattctaatatgtggggacattaaaaattctcgt
 gtcgcaagaagtaattatcatagtttaacatcattaggtgccaacgtaattgttctcaagt
 ccaaaagaatgggttagataatacattagaggcgcttatgttgaaattgatgaagtcatt
 gataaagtagatattgttatgttgcttagagttcaacatgaaagacatggaatttcagggt
 gaagctaactttgctgctgaagaatatcatcaacaatttggtttaacacaggctagatat
 45 gataaattaaaagagggaagccattgtaatgcacagctcctgtaaatagaggtgttgaa
 attaaaagcgagctagttgaagcacctaagtcctogaatatttaagcagatggaaaatgga
 atgtatttaagaatggcagtaataagtgcgcttttacaatag

Sequence 2602

50 MEHLLSMEHLSNSEIYDLITIAQCFKSGERPLPQFNGQYVSNLFFENSTRTKCSFEMAEQ
 KLGLKLINFETSTSSVKKGESLYDTCKTLESIGVDLLVIRHSQNSYYEELDQLNIPIANA
 GDGSGQHPTQSLLDIMTIYEEYGSFEGLNILICGDIKNSRVARSNYHSLTSLGANVMFSS
 PKEWVNTLEAPYVEIDEVIDKVDIVMLLRVQHERHGISGEANFAAEEYHQQFGLTQARY
 DKLKEEAIVMHPAPVNRGVEIKSELVEAPKSRIFKQ MengMYLRMAVISALLQ*

55

Sequence 2603

Contig_0778_pos_3972_5249,
 is similar to (with p-value 0.0e+00)
 >sp:sp|P46538|PYRC_BACCL DIHYDROOROTASE (EC 3.5.2.3) (DHOASE

). >pir:pir|S34319|S34319 dihydroorotase (EC 3.5.2.3) - *Bacillus caldolyticus* >gp:gp|X73308|BCPYR_2 *B. caldolyticus* pyrimidine biosynthesis genes. NID: g312439.
 atgaaattaattaaaaacggaaaaatcttaaaaaacgggtatcctaaaagacacagaaatt
 5 ttaatcgacggtaaacgtattaaacaaattagtagtaaaattaatgcttcattcttcaaat
 attgaagtattgatgcaaaaggaaatttaattgctcccggttttgtagatgttcattgtg
 cacctacgtgaaccaggtggtgaacataaagaacaattgaaagtggtaaaaagccgct
 gcaagaggtggttttactacagtatgtcctatgcctaatacaagacctgtaccagataca
 gttgaacatggttagagaattaagacaacgaatttctgaaacagcacaagttagggtgttg
 10 ccttatgctgctattactaagagacaagcaggtactgaacttggtgattttgaaaaatta
 gcactagaaggtgtgtttgcatttactgacgatggtgtgggagttcaaacagcaagtatg
 atgtatgctgctatgaagcaagctgcaaaagttaaaaaacgattgtcgcacactgtgaa
 gataatagcttaattctatggtggtgcaatgcataaaggtaaacgtagtgaaagaattaggc
 atacctggtattccaaatattgctgaatctgtacaaattgctagagatgtattattggct
 15 gaagcaactggttgtcactatcatgtgtgtcatgtttcaactaaggaaagtgttcgagta
 atcagagacgctaaaaaagctggtatccatgtaacagcagaagttacaccacatcattta
 ttatttaactgaaaatgatgttcctggcgatgattcaaaactacaaaatgaatccaccatta
 agaagtaatgaagatagagaagcacttttagaaggcttattagatggaacaattgattgt
 attgcaacggatcatgcacctcacgctaagaagaaaaagcacaacctatgacaaaagca
 20 cctttcgcatcgtaggtagtgaaacagcattccattactttatacacactttgtgaaga
 cgaggttaattggtcactgcaacaattagttgattatttctactattaaccagctattatt
 ttcaacttaattatggaaaattacacaaagatagttacgctgatttaacaataattgat
 ctttaactgaaaaagaaatcaaaagtgaagatttcttatctaaagctgataacactcca
 tttattggtgaaaaagtttatggaaatccaacactaacaatgcttaaggtgaagtagta
 25 ttcgaggaggaaaaagtag

Sequence 2604
 MKLIKNGKILKNGILKDTTEILIDGKRIKQISSKINASSSNIEVIDAKGNLIAPGFVDVHV
 HLREPGGEHKETIESGTKAAARGGFTTVCPMPNTRPVPDTEHVRELQRRISETAQVRVL
 30 PYAAITKRQAGTELVDPEKLALEGVFAFTDDGVGVQTASMMYAAMKQAAKVKKPIVAHCE
 DNSLIYGGAMHKGKRSEELGIPGIPNIAESVQIARDVLLAEATGCHYHVCHVSTKESVRV
 IRDAKKAGIHVTAEVTPHLLLTENDVPGDDSNYKMNPPLRSNEDREALLEGLLDGTIDC
 IATDHAPHAKEEKAQPMTKAPFGIVGSETAFPLLYTHFVRRGNWSLQQLVDYFTIKPATI
 FNLNYGLKHKDSYADLTIIDLNTEKEIKSEDFLSKADNTPFIGEKVYGNPTLTMLKGEVV
 35 FEEK*

Sequence 2605
 Contig_0778_pos_5250_0,
 is similar to (with p-value 5.0e-51)
 40 >sp:sp|P25993|CARA_BACSU CARBAMOYL-PHOSPHATE SYNTHASE, PYRIMIDINE-SPECIFIC, SMALL CHAIN (EC 6.3.5.5) (CARBAMOYL-PHOSPHATE SYNTHETASE GLUTAMINE CHAIN). >pir:pir|E39845|E39845 carbamoyl-phosphate synthase (glutamine-hydrolyzing) (EC 6.3.5.5), pyrimidine-repressible, small chain - *Bacillus subtilis* >gp:gp|M59757|BACPYROP_5 *Bacillus subtilis* pyrimidine biosynthetic (pyr) gene cluster (pyrR, pyrP, pyrB, pyrC, pyrAA, pyrAB, pyrD, pyrF and pyrE) genes, complete cds. NID: g387576. >gp:gp|Z99112|BSUB0009_22 *Bacillus subtilis* complete genome (section 9 of 21): from 1598421 to 1807200. NID: g2633902.
 50 atgcttgaaaaacgttatcttgactggaagatggctcttattacgaaggatcgcgtta
 gggtaagatgacttatctataggcgaaattgtattcaacactgctatgacgggtaccaa
 gaaacaatctctgaccgctcttacacaggtcaaatcataacttttacgtaccactaatt
 ggaaactatggtattaatcgcgatgattttgaatcattaacacctaaattaaatgggga
 55 gtagtaaaaggaagcaagtacacaccctagtaacttttagacacacaaaaaactttacacgaa
 acacttgctcaatatcatattcctggtatatcggggtagataactagaagtattactcgt
 aaaattagaaattatggtgttttaagagctggatttacagataataaagataaacattcag
 gaacttggtgaacagttgaaaactgctgaattacctagagatgaagttcaaacggtttct
 acaaaaacacccatattgatcaacaggttccgatttaagcgtcggttttactcgactttggt
 aaaaagc

Sequence 2606

MLEKRYLVLEDGSYYEGYRLGSDDLISIGEIVFNTAMTGYQETISDPSYTGQIITFTYPLI
 GNYGINRDDFESLTPKLNQVVVKEASTHPSNFRHQKTLHETLAQYHIPGISGVDTRSITR
 5 KIRNYGVLRAAGFTDNKDNIQELVEQLKTAELPRDEVQTVSTKTPYVSTGSDLSVVLDFG
 KKK

Sequence 2607

Contig_0779_pos_3369_1672,
 10 is similar to (with p-value 0.0e+00)
 >sp:sp|P18255|SYT1_BACSU THREONYL-TRNA SYNTHETASE 1 (EC 6.1.
 1.3) (THREONINE--TRNA LIGASE) (THRRS). >pir:pir|B37770|YSBST
 1 threonine--trna ligase (EC 6.1.1.3) 1 - Bacillus subtilis
 >gp:gp|AF008220|AF008220_195 Bacillus subtilis rrnB-dnaB gen
 15 omic region. NID: g2293135. >gp:gp|M36594|BACTRNAB_1 B.subt
 ilis threonyl-trna synthetase (thrSv) gene, complete cds. NI
 D: g143765. >gp:gp|Z99118|BSUB0015_160 Bacillus subtilis com
 plete genome (section 15 of 21): from 2795131 to 3013540. NI
 D: g2635200. >gp:gp|Z75208|BSZ75208_5 B.subtilis genomic seq
 20 uence 89009bp. NID: g1769994.
 atggcacaagcattaaaaacgtttatacggagacgttaaatttggagttggacctgtaata
 gaaggcggattctattatgattttgataggatgataaggtttcatcgatgattttgat
 aaaattgagaaaaaatgaaacaaattgtgaacgaaaatcataaaattgtaagagaagta
 gttagtaaaagaaaaagcaaaagacttcttcaaggatgaccttataaaattagaacttatt
 25 gatgcaattcctgaagatgagagtgtaaacactttataactcaagggtgaatttactgattta
 tgtcgaggtgtacacgtaccttctacttctaaaaattaaagagttcaaaactattatctaca
 gctgtgcttatttggcgtggaaatagtataataaaatgttacaacgaatttatggtaca
 gcattctttgacaaaaaagatttgaaagcacatctaaaaatgttgaagaacgtcgtag
 cgtgatcatcgtaaaatttgtaagatttagaattgtttacaacaatcaactcggtggt
 30 gctggtttaccattatggttaccaaatggtgctacaatacgtagggaaatagaacggtat
 attgtcgataaagaagtaagtatgggatacgtatcatgtttacacaccagtattagccaat
 gttagattatataaaacatctggtcactgggatcattatcaagaagatatgttccagca
 atgaagttagatgaagacgaagcaatggtcttaagaccaatgaactgtccacatcatatg
 atgatttataaaaaacaaacctcattcttctcggaattacctatacgtattgctgaattg
 35 ggtactatgcacgtttacgaagcaagtggtgcagtatcagggtttacaacgtgttcgagga
 atgacattgaatgattcccatattttcgttagacctgatcaaattaaagaagaatttaaa
 cgtgtagttaatatgattcaagatgtgtacaaaagattttggttttgaagattatcgcttc
 agattgagttatagagatcctgaagataagcataagtaactttgatgatgatgaaatgtgg
 gaaaaagctgaatccatgcttaagaagcatcagatgaattaggttttaacttatgaagaa
 40 gctattggtgaggcagcatctatggacctaaagtttagatgttcaagtaaaaaacagctatg
 ggaaaaagaagaactctatcaacagcacaaacttgattttctttaccagaacggtttgac
 ttaacgtacattggtcaagatggagaacaacatcgctcctgtagttatacaccgtggtgta
 gtttctactatggaacggtttgttgcatttttaacagaagaacaaaagggtgcatttcca
 acttggttggcgctatgcaagttgaaattattcctgtaaatatagatttacattatgat
 45 tatgcaagacttttacaagatgaactaaaatcgcaaggtgtccgcgttgaaattgatgac
 cgtaatgaaaaaatggatataaaattcgtgaagctcaaataaaaaaaccttatcag
 attgttgttaggtgaccaagaagtagagaatcaagaagtaaatgtaagaaaatatggttct
 gaaaaacaagaatcagttgaaaaagatgaatttatttggaaatgttattgatgaaatccgt
 ttgaaaaagcatagataa

Sequence 2608

MAQALKR.LYGDVVFVGVPIEGGFYYDFDMDKVVSSDDFDKIEKTMKQIVNENHKT.VREV
 VSKEKAKOFFKDDPYKLELIDAIPEDSVTLTYTQGEFTDLCRGVHVPSTSKIKEFKLLST
 AGAYWRGNSDNKMLQRIYGTAFDCKDLKAHLKMLEERRERDRHKIGKDLELFTNNQLVG
 55 AGLPLWLPNGATIRREIERIYVDKEVSMGYDHVYTPVLANVDLYKTSGHWDHYQEDMFPA
 MKLDEDEAMVLRPMNCPHHMMIYKNKPHSYRELPIRIAEELGTMHRYEASGAVSGLQVRVG
 MTLNDSHIFVRPDQIKKEEFKRVVNMIQDVYKDFGFEDYRFLSYRDPEDKHXYFDDDEM
 EKAESMLKEASDELGLTYEEAIGEAIFYGPKLDVQVKTAMGKEETLSTAQLDFLLPERFD
 LTYIGQDGEQHRPVVIHRCVVSTMERFVAFLEETKGAFPTWLAPMQVEIIPVNIDLHYD

YARLLQDELKSQGV RVEIDDRNEKMGYKIREAQMKKIPYQIVVGDQEVENQEVNVRKYGS
EKQESVEKDEFIWNVIDEIRLKKHR*

Sequence 2609

5 Contig_0781_pos_5154_4732,
is similar to (with p-value 6.0e-49)
>gp:gp|D88209|D88209_1 Bacillus licheniformis DNA for Pz-pep-
tidase, complete cds. NID: g1651215.
atgtttgctgaatttgaacataaaatcacatcaaataagaagctggggagccggttaacg
10 ccaaatcgatgaatgaagaatatgctaaactgaacaaactatatatttgggtgaagcagta
gaaactgacgatgatattagtaaagaatggtcacgtattcctcatttctatatgaattat
tatgtatatcaatacgcgaactggttatagtcagctcaaagtttaagtcacaaatttta
actgaggggtcaacctgctgttgaacgatataatgaattcttaaaaaagggtagctca
aactatccgattgaaattttaaaaaatgcaggtgttgacatgacaacacctcaaccaata
15 gaggaagcttgtgaagtattcgaacaaaaattagatgcttttgaaggttaataaagct
tag

Sequence 2610

20 MFAEFEHKIHQIEEAGEPLTPNRMNEEYAKLNKLYFGEAVETDDDISKEWSRI PHFYMN
YVYQYATGYSAAQSLSHQILTEGQPAVERYINEFLKKGSSNYPIEILKNAGVDMTTPQPI
EEACEVFQKLDFAFEKLMKA*

Sequence 2611

25 Contig_0781_pos_3771_3247,
putative peptide of unknown function
atgtatgttggtacgagtaaaagtagtggttttaaatctaaactgcagttcttaaggttaa
gatatttatcatgtatatgctgaatataagctctccatattataagcaatatggtaaatca
gaagcccctacaatatatgatgataatattacaagtcaatcagagttaaagaagaatta
aaagaacacttgatgacataccaacaatcgaagtagcaacgaattatttaggattagaa
30 agtattcatgaaaataatactattcgatttatacacaacctatcgatttaatactgat
ttaaaagttgtcaaaacttactgaatatcacccttgtttcgcagcctattgaagtggaa
ttcagtaatgcacagaaagatattataaaaaatgcaatcacagttcaatcgtaggttaaga
aaggttaataatcttatgaaaaaaggattcaaaactagtactattctttaaagtgttta
gaggaatataacgaaacagtaggaagtgtattgattgatgagtaa

Sequence 2612

40 MYVGTSSKSSVLNLTAVLKGKDIYHVYAEYKSPYKQYKGSEAPTIYDDNITSQSELKKKL
KETLDDIPTIEVATNYLGLSEIHENNTIRFIHKPIGFNTDLKVVKLTEYHPLVSQPIEVE
FSNAQKDIKMQSQFNRRRLRKVNLMKKGFKTSYSLNVLEEYNETVGSVLIDE*

Sequence 2613

Contig_0781_pos_3065_1200,
putative peptide of unknown function
atgttattaacttttagactttcctatttcaaataggacacacatttagaaccaagatgata
45 aataatttttagaacaatacttaattattataatgaattagatcatcagcatcgcgacac
acagaaactaagcatcatgcacatcaagccatgcaggttgattatagaaatacaaacggt
tctgcatttttagattatcttaacggtaatatgaatgggcttggttttaggagcaaatgga
gacggtatagctgaaacaaaacaagccagagtatcaatagatggtaccgtacatcccttg
ttgcaagaaaggctgcttcatgacttttttaggaattaacagaaaattagataaagaaata
cattctaattggtgcagttgattttatttggaaatcctccatataaccaggaaatagattg
50 ggagaaaatgggacaccaataattgggaaccagaagcccatattgaagcgtttttaaac
cctttagttgataatcaatacgttacaaaagaagttataggagaagatacatcaggaaaa
tataatgtgtacaaatttacgtttgaaccacaaaattacaataaaacgttacttattact
tcattgatacacggtaatagaactactggattttttgatatgtgccatatactcaatcta
15 ttggtcaatcaattgggaaaagtatcctcaattaacttacttaagaaaaaatgtacgttta
atattatgttctatggttaaccggtggggttcgcaaatcaagaaagagagaatgtgaac
aatgtagattttaaacagaaattttgattataactggaaggcaggttaagggacagatcct
gataaatctaacttcaaaaggtaaaagtccttttctgaaaaagaatcacaaaatatgcgt
agcttagttcaagtatagataatttaactgctcacttagatttgcatgatattatttca

gtaaataatgattactgtttatcttccgcgttgggccaatcaaaaaataataatg
 actcatcttattaacaatttaaaaagtaacggagacctcgttgttgggggtccagtaca
 ttatcatcttttagtaattgggttaggaatccgaaataaaacaacgtcatatctttcagaa
 ataaatgaaaaacgtgtcgttgaaaagaaaagtcggaagaaatgagacgttcagtacgc
 5 tgggttaggtaatgtaatttttagaatggcacaatttgaatcttatcaaaatggcacaaca
 tcattagatcctttcattaaagtgtatgatgatagatttaacaataaaacatct
 gaagtcatcaccctacgtgcagaaaggaatgaatggcaacgtataatgatgagtcagcag
 cgtttcaaaagtttttagcaaatggattttagagctctatggatatgtgactataaacgtt
 gatagagatgtcacagtggtggttaatacctaatttggcagaattatcatccattcttt
 10 ggatttaataaaagtagaaaacgtaatttttcaattgaacatagactcaacaaagga
 aatacaactttccctatttacgctgctgctggagttcaaatgtcgacgattactgaacca
 ggtacaaaacgtactgatacagtaatgccggtactagatgtaagaaaaaagggtgctggt
 attgtaacaatacaacaaattaaattatttgcgaagttcactcctacgcattctgctaatt
 tccattcagatattaaaaatctggagaatatttttttttttctaacgctttgcat
 15 atggttcattatccacatttctgaatttgtggttctttattcggtaagccctctc
 tcatag

Sequence 2614

MLLTLDFPIQIGHTFRTKMINNFRTILNYYNELDHQHRAHTETKHHAHQAMQVDYRNTNV
 20 SAFLDYLNNGNGLVLGANGDGAETKQARVSIDGTVHPLLQERLLHDFLGINRKLDEI
 HSNAGVDFIWNPPYIPGNRLGNGTPNNWEPEAHIEAFLNPLVDNQYVTKEVIGEDTSGK
 YNVYKFTFEPQNYNKTLLITSCIHGNETTGFFDMCHILNLLVNQWEKYPQLTYLRKNVRL
 IYVPMVNPWGFANQERENNVNLDNRNFDYNWKAGKGTDPDKSNFKGKSPFSEKESQNM
 25 SLVQSIDNLTAHLDLHDIISVNNDYCLFYPRWANQKNNMTHLINLKSNGDLVWGSST
 LSSFSNWVGIRNKTTSYLSEINEKRVGEKKSPEEMRRSVRWGVNIFRMAQFESYQNGQT
 SLDPFIKVMVYDDRFNNKTSEVITLRAERNEWQRIMMSQQRKVLANGFVELYGYVTINV
 DRDVTVGINPNIVQNYHPFFGFENKSRKRNLFIEHRLNKGNTTFPIYAAAGVQMSTITEP
 GTRKTDVMPVLDVKKKGAGIVTIKQIKLFAKETPTHSANSIQILKSGEYFYFFLTLCD
 30 IAHHYPTFLEFVVLVSVSHLS*

Sequence 2615

Contig_0784_pos_1894_2202,
 putative peptide of unknown function
 gtactaaagcttctaagattttgtagtaaacattaacatatgtcctccgaatagatct
 35 tggatagcatttgaatcagtaactccaggtacaatcggtacgactgcaatgataatt
 gtagctaaatctccgctaggaacaaatgcatgtccaattacagaaataatacctattacc
 aaagaacctatgaattctggaataaattgtgcgtgtagcttctgatctaatttctact
 actaagtatccaatcggtccagctaatacagctgtgatgatacaaccagacgacctccc
 40 tgtagatag

Sequence 2616

VTKASKDFVNNINICPPNRSWIAFVISTPGTIGMTAAMIIVAKSPLGTNACPITEIIPIT
 45 KEPMNSGINCACSFRSNISTTKYPIVPANTAVMISTRPPCR*

Sequence 2617

Contig_0784_pos_4022_5011,
 is similar to (with p-value 0.0e+00)
 >pir:pir|S39743|S39743 hypothetical protein - Bacillus subti
 lis
 50 atggctgatttattatctgtattacaagacaaattatccgggaaaaatgtaaaaatagta
 ttacctgaaggtgaagatgaacgagtgctcattgctgcgactcagctacaaaaaactgac
 tatgtttcacctatcggttctaggaaacgaagataatattaaatctcttgccttctaacaac
 gctttgattttaactcaaattgaaatcatagatccagcaacgagtgaaacttaagatgag
 55 ctgtagatgcttttggtaagacgtaaaaggtaaggcaactaaagaacaagcagtgaa
 ttattagataatgtaattatttccggaacaatgcttgtgtatactggaaaggctgagggt
 ttagtgagtggtgctgcacattctactggagatacagtcagaccagcattacaaattatc
 aaaactaaacctggtgtatctagaacatctggtattttctttatgattaaaggcgacgaa
 caatatattttggagattgtgcgattaatccagaattagatgctcaaggacttgctgaa
 attgcagtagagagtgctaaatcagcacaaagctttggaatggaccctaaagtagctatg

5 ttaagcttttctacaaaaggttctgctaaatcggtatgatgttactaaagtgaagaagca
 ttgaagtttagctcaagaaaaagctgaagcagatcaattagatcatgtagttattgatgga
 gaattccaatttgacgctgctattgttcctagcgtagcagagaagaaagcacctgggtgca
 aaaattcaaggtgatgcaaatgtattgggtttccctagtctagaagcaggtaatattggt
 tataagattgctcaacgttttaggtggatacgatgcagtaggaccagtcctacaaggatta
 aactctccagtcgaatgatttatctcgtggttgcactgaagacgtttataacttatct
 attattacagctgctcaagctttacaataa

Sequence 2618

10 MADLLSVLQDKLSGKNVKIVLPEGEDERVLIATQLQKTDYVSPIVLGNEDNIKSLASKH
 ALDLTQIEIIDPATSELKDELVDVAFVRRKKGKATKEQAVELLDNVNFGTMLVYTGAEG
 LVSGAAHSTGDTVRPALQIIKTKPGVSRSTSGIFFMIKGDEQYIFGDCAINPELDAQGLAE
 IAVESAKSAQSFGMDPKVAMLSFSTKGSAKSDDVTKVQEALKLAQEKAQADQLDHVVIDG
 EFQFDAAIVPSVAEKKAPGAKIQGDANVLVFPSEAGNIGYKIAQRLGGYDAVGPVLQGL
 15 NSPVNDLSRGCTEDVYNLSIITAAQALQ*

Sequence 2619

Contig_0784_pos_5083_5850,

is similar to (with p-value 5.0e-66)

20 >sp:sp|P39648|YWFL_BACSU HYPOTHETICAL 31.4 KD PROTEIN IN PTA
 3'REGION. >pir:pir|S39745|S39745 hypothetical protein - Bac
 illus subtilis >gp:gp|X73124|BSGENR_91 B.subtilis genomic re
 gion (325 to 333). NID: g413923. >gp:gp|Z99123|BSUB0020_60 B
 acillus subtilis complete genome (section 20 of 21): from 37
 25 98401 to 4010550. NID: g2636240.
 atgcaatcttttgcgttgatgacactttttccgaaagcgttggtaaagatttatcttgt
 aatgtagtacgaacgtggatacatcaacacaccgtgattttgggcattcatgattcgcgt
 ttaccatttttaagtgtggtattcgtttttctacagatgaacaaggatataatgcaatt
 gtttaggaattctggtggcttggtgctgatttagatcaaggaattttaacatatctttg
 30 atttttaaggacaaaccgaaacgactattgatgaagcctttacagtgtgtatttattg
 attaataaaatgtttgaggatgaagatgtagtatcgatactaaagaaattgagcaatcg
 tattgcccgagaaaatttgatttaagtattaatgataagaaatttgccgggatttcgcag
 cgacgagtacgtggtggtatcgacgtgcaaatatacttatgtattgaaggttctggctca
 gaacgggcatttaattgatgcaacagttttatcaacgtgcttaaaagggagactactaaa
 35 tttcactatccagacatagatccctcatgtatggcatcttttagaaacccttttaataga
 gaaattaaagtgaagatgttatgtttttattattatgcactaaaagatttaggggca
 aacttaaatatggatcctattacagaagacgagtggaacggttacgaagggtattatgat
 aagatgttagaacgcaatgcgaaaatgaatgaaaaattagatttttag

Sequence 2620

40 MQSFAFDDTFSESVGKDLSCNVVRTWIHQHTVILGIHDSRLPFLSDGIRFLTDEQGYNAI
 VRNSGGLGVVLDQGIILNISLIFKGQTETTIDEAFTVMYLLINKMFEDEDVSI DTKEIEQS
 YCPGKFDLSINDKKFAGISQRRVRGGIAVQIYLCIEGSGSERALMMQQFYQRALKGETTK
 FHYPDIDPSCMASLETLLNREIKVQDVMFLLLYALKDLGANLNMDPITEDEWTRYEGYYD
 45 KMLERNAKMNEKLD*

Sequence 2621

Contig_0784_pos_3890_3102,

is similar to (with p-value 9.0e-94)

50 >sp:sp|P39645|YWF1_BACSU HYPOTHETICAL 29.5 KD PROTEIN IN ROC
 C-PTA INTERGENIC REGION. >pir:pir|S39742|S39742 hypothetical
 protein - Bacillus subtilis >gp:gp|X73124|BSGENR_88 B.subti
 lis genomic region (325 to 333). NID: g413923. >gp:gp|Z99123
 |BSUB0020_63 Bacillus subtilis complete genome (section 20 o
 55 f 21): from 3798401 to 4010550. NID: g2636240.
 atgaatccaataaaaaatttgaaaaaggagcgaataataatgagtgaagcagcagaaact
 ttgatgggttggtatagcttacattttattttatgcagtagactggacaacttttcgttta
 attgctgaagatgatcgtgaagcaatgattactgaattgaaacattttattaaagataaa
 acagttgctagagaatcacatcaaggtgatcatgcaattttataacattacaggtcaaaaa

gcggaccttttactatggtttttacgtccagaaatgaaagagttaaatcaaattgaaaat
 gagtttaataaattacgtatcgagactatctcattccaacttattcctatgtgtcagt
 atagaattaagtaattattagcaggcaaatctgatgaggatccttatgaaaatccccac
 gtaaggcacgattataccctgaattaccacattctgaatatatatgtttctatccaatg
 5 gataaacgacgcaatgaaacttataactgggtatgttacctatcgaagaccgtaaaact
 ttaatgtataaccatgggatgtaggtcgtaaatatgctggtaaaatcaaacagtttatt
 acaggttcagtaggttttgatgactatgagtggggtgttacattattttcaaagtatgta
 cttcaattcaaaaaaattgtctatgaaatgcgttttgatgaaacgactgctcgttatggc
 gaatttggtagtttctatattggtcacattctaaacatcgaagacttcaacaatttttt
 10 agtatataa

Sequence 2622
 MNPIKILKKERIIMSEAAETLDGWYSLHLFYAVDWTFRLLIAEDDREAMITELETFIKDK
 TVARESHQGDHAIYNITGQKADLLWFLRPEMKELNQEENFNKLRIADYLIPTYSYVSV
 15 IELSNYLAGKSDDEPYENPHVKARLYPELPHSEYICFYPMDKRRNETYNWYMLPIEDRKT
 LMYNHGMIGRKYAGKIKQFITGSGVFDYEWGVTLSNDVLQFKKIVYEMRFDETTARYG
 EFGSFYIGHILNIEDFKQFFSI*

Sequence 2623
 20 Contig_0784_pos_2495_1845,
 putative peptide of unknown function
 atggcacgtattgctacaaaattgggctatcctgaaagtaatagtttcgtgactaatact
 gtaattgaatttgttttacataatgaagcatatcctcggttgatagaattaaaactcga
 gatacgaacttaataaaaaatttctcaagctaataatctcacgtcaaattacaaatggc
 25 acaatgacgcttgaagaagctaagtatcaattagaggaaatatatgttgctaaaagagat
 agcagttctcccttttaaggaattgccgcagcaattatcgctacgagcttccctctatcta
 cagggaggtcgctggttgatcatcacagctgtattagctggaacgattggatactta
 gtagtagaaatattagatcgaaagctacacgcacaatttattccagaattcataggttct
 ttggtaatataggtattatttctgtaattggacatgcatttgttcctagcggagatttagct
 30 acaattatcattgcagcggctcatgccgattgtacctggagtactgattacaaatgctatc
 caagatctattcggaggacatatgttaattgtttactacaaaatctttagaagcttttagtc
 accgccttttggtataggcgctggtgtaagttcaatattaatttttagtctag

Sequence 2624
 35 MARIATKLGYPESNSFVTNTVIEFVLHNEAYPRLYRIKTRDTNLIKISQANEISRQITNG
 TMTLEEAKYQLEEIYVAKRDSSLPFKGIAAAIIATSFLYLQGGRLVDIITAVLAGTIGYL
 VVFILDRKLHAQFIPEFIGSLVIGIISVIGHAFVPSGDLATIIIAVMPIVPGVLITNAI
 QDLFGGFMLMFTTKSLEALVTAFIGAGVSSILILV*

Sequence 2625
 40 Contig_0786_pos_2219_1125,
 is similar to (with p-value 1.0e-53)
 >sp:sp|P39604|YWCF_BACSU HYPOTHETICAL 43.3 KD PROTEIN IN QOX
 D-VPR INTERGENIC REGION. >pir:pir|S39697|S39697 hypothetical
 45 protein - Bacillus subtilis >gp:gp|X73124|BSGENR_43 B.subtilis
 genomic region (325 to 333). NID: g413923. >gp:gp|Z99123
 |BSUB0020_107 Bacillus subtilis complete genome (section 20
 of 21): from 3798401 to 4010550. NID: g2636240.
 atgggtggcggacaatatagtgcaaatcttcagtattagacaaattatttattatattttc
 50 ggagctattattgcatttttaattatgataatttcaccgaaaaaaattaaaaataatact
 tatattttatacagtatatatttgcgttctattaatagggttacttattttacctgaaact
 tcaatcaccccaattattaatggtgctaaaagttggtacagtttcggtcctataagcatc
 caaccttcgaattcatgaaaattatacttatacttgctttggctaaaacgatatctaaa
 cataaccaatttacatttaataagtccttttcagctctgatttaattgttattttttaaatt
 55 ttaggtgtatccattataccatggtcatttaattctattgcaaaatgacctaggtactact
 ttgggtgttatgtgcaattatagctggcgctcatgttagtaagtggaataacatggaggata
 ttggccctctttttattgttgcatttgtaagtgggttctagtattatattagctatcatt
 tataaacatccttaatagaaaacctattaggaataaaaatgtatcaaatgggacgtatc
 aattcttggttagatccctattcatacagtagtgagatggatataccttaacagaatct

ttaaaagctattgggttcagggtcaattatttaggttaaagggtataaccatggcgaagtttat
 attcctgagaatcataccgactttattttttcagtgattggagaagaaatgggctttata
 gggttcagttattattgatattactttttcttattcttaatatcccttatacggtttagct
 agtaaaattgatagtcagtttaacaaagtatttatcataggatatgtatcgttgattggt
 5 tttcacgtgttacaaaatatcggcattgacgggtcaattattaccgattacaggtatacca
 cttccgtttatttagttacgggtggaagttctttatggagtttaattgactggatagggagta
 gttctttcaatttattatcatgaaccccaagatatgaaataaccacattatctaaaaaa
 tctaatacaatttaa

10 Sequence 2626
 MGGGQYSANFSIRQIIYYIFGAIIFLIMIISPKKIKNNTYILYSIFCVLLIGLLILPET
 SITPIINGAKSWYSEFGPISIQPSEFMKIILILALAKTISKHNQFTFNKSFQSDMLFFKI
 LGVSIIPMALILLQNDLGTTLVLCIIAGVMLVSGITWRILAPLFIVAFVSGSSIILAI
 YKPSLIENLLGIKMYQMGRINSWLDOPYSSSGDGYHLTESLKAIGSGQLLGKGYNHGEVY
 15 IPENHTDFIFSVEEMGFIGSVLLILLFLIFLIRLASKIDSQFNKVFIIGYVSLIV
 FHVLRNIGMTVQLLPITGIPLPFISYGGSSLSLMTGIGVLSIYYHEPQRYEITTLSSK
 SNTI*

Sequence 2627
 20 Contig_0787_pos_1817_2143,
 putative peptide of unknown function
 atgattttctttatcattaactttattcttatttcatatttgtacgataaacaatatgta
 ccttttcaagcaattaccggtataagcttgtttattgcttagttatatttccaataaca
 ctcattttatattgtgcgtattgccaaaaaaattatctatacagtaataagtatgaaatg
 25 agaactggaataatcattggtattattgctttaaattctagtaattatgcaagggtttcac
 ttttaactgggctatagatttttttaataatgttgtatggtggtcatttaaaagtaccatc
 tcaattgatagacgaatcgcaacatag

Sequence 2628
 30 MIFFIINFILISYLYDKQYVPFQAITGISLFYCLVIFPITLILYVRIAKKNYLYSNKYEM
 RTGIIIGIILVIMQGFHFNWAIDFFNNVWWSFKSTISIDRRAT*

Sequence 2629
 35 Contig_0787_pos_6394_6038,
 putative peptide of unknown function
 atgccagaaaatattgtaagaacaaaagggttagtgatggtagcgcagtataatgatgta
 gcgtgtttgttatcacaggctgggtcatcttgaatattcaccggttacatactgggtg
 gcaacaatgagtgaaagtcaacagcaagctattttggaggcgcgtcaagatgtagtagaa
 gattgggatatcgaatatggagatcgtcaaacgcaatttgaattattggtacggattta
 40 gatcaagaaaaaattttccgggaattagatgcatgcttaattcatagtagtgagattgat
 gaagattggcgattactagatagtcgcgtatcaatggacttatgatcgacgaatgtaa

Sequence 2630
 45 MPENIVRTKGIVWLAQYNDVACLLSQAGSSCNIHPVTYVWVATMSSESQQQAILEARQDVVE
 DWDIEYGDRTQFVIIGTDLDQEKISRELDACLIHSSEIDEDWRLLDSPYQWTYDRRM*

Sequence 2631
 Contig_0787_pos_4438_2969,
 is similar to (with p-value 0.0e+00)
 50 >sp:sp|P39755|NDHF_BACSU NADH DEHYDROGENASE SUBUNIT 5 (EC 1.
 6.5.3) (NADH-UBIQUINONE OXIDOREDUCTASE CHAIN 5). >gp:gp|U283
 23|BSU28323_1 Bacillus subtilis NADH dehydrogenase subunit 5
 (ndhF) gene, complete cds. NID: g903586. >gp:gp|Z99104|BSUB
 0001_183 Bacillus subtilis complete genome (section 1 of 21)
 55 : from 1 to 213080. NID: g2632267. >gp:gp|Z99105|BSUB0002_11
 Bacillus subtilis complete genome (section 2 of 21): from 1
 94651 to 415810. NID: g2632457. >gp:gp|AB006424|AB006424_9 B
 acillus subtilis genomic DNA, 70 kb region between 17 and 23
 degree. NID: g3599592.

atgttttttattacacttgcattgctatactgagtgaggattaatatttttgaatcctcga
 gttcccatccaataatataaaatttcatataatttacttgttttacctatcattacggga
 ttaagtggattaatatttttcggtgaaagggcgaatgttggacctttttagttgatcat
 cttacttggtaatgatgacatttattttgactttaggctttatcattcaaaagttttct
 5 atgcgttatttaattggcgacatgcattaccgtaaatattttccgttttttaccattaatt
 actgcatttgccttcattggcatgggtaagtggagacttaaggttaatgaccatgttttgg
 ggtgcaacattatttgtgttaacacggctcattaaagttaacaaattatggaagggtgcct
 aggggaagcagcaagaatttcagcttgggtcatttatattggcatggttgcgttattgatt
 gctgtcattttattgtatatcgctacaggagattgggtatatttattcgaatatgtcagat
 10 gataatgcaatcaattatggaatgcgtctctgtatcaatttacttattgttttagctgtg
 attattccggcggcacaatttccatttcaaggctggcttattgaatctgtagctgcgcct
 acggcagtttcagctattatgcacgctgggtattgttaatgctggtggcggtattcttaca
 cgcttttctccggtatttaatgacgaaatagccatttctactgttattaattattgcaagt
 atttcagttattgttgggttctggaatcagctctgtgcatgttgattacaagagacaactt
 15 gtaggttctacgataagtcaaatgggttttatgttagtacaatgtgcgcttggggcatat
 tctgcggcgatagtagacatttaatttgcagtggtgtgtttaaagcgacattatttctaca
 tcgggttctgttgtttaaagatttaacattcctacgcctccatctgttaaaaaatcatat
 ggctggcttgtatttggctcgtctactagctattcttatagcgataatattttggttgaat
 agtgatagacatgcatatgatgtattaaagcgtctcttatattagcttggctcgttaattggtg
 20 tcttqgaatcaattagttgcttttagtcatggactcatcggaagagtcacggagtatgt
 atgactattgtttagtaattgtttatattattacgcacattatttcttcacgacatta
 agtaacgttgatattcatattgtttcaccaccactcataagtattattctatcgattgct
 attatagttttcggcagtatgttaagcatatgggtatcacggcgaagagaatcaaaggca
 tttgcgaagttatacttgtggcttatttaaagtaggagaggctaaaaccaatctatagaa
 25 agtcatccatcatattttaaaccgatttttag

Sequence 2632

MFFITLVIAILSGLIFLNHRVPIQYIKFHIYLLVLPITGLSGLIFFGERANVGPFVVDH
 LTWLMFTFILTLGFIIQKFSMRYLIGDMHYRKYPFFTLITAFASLAWLSGDLRLMTMFW
 30 GATLFVLTRLIKVKNLWKVPREARISAWSFILAWLSLLIAVILLYIATGDWYIYSNMSD
 DNAINYGMRLCINLLIVLAVIIPAAQFPFQGWLIESVAAPTVPVSAIMHAGIVNAGGVILT
 RFSPVFNDIAISLLIIASISVLLGSGISLVHVDYKRQLVGSTISQMGFMLVQCALGAY
 SAAIVHLILHGVFKATLFLQSGSVVKRFNIPPPSVKKSYGWLVFGRLLAILIAIIFWLN
 SDRHAYDVLALILAWSLMVSWNQLVAFSHGLIGRVIGVCMIIVVVIVYIITHHYFFTTL
 35 SNVDIHIVSPPLISIIISIAIIVFGSMLSIVVSRRESKAFKLYLWLKIVGEAKTQSIE
 SHPSYLKRF*

Sequence 2633

Contig_0787_pos_2951_2058,

40 putative peptide of unknown function

atgcitcaatcagatatcaatgaattagtcattcaggctaaaacgtgtaattacaccttta
 tcaccgatttcaacatttgcgtgcccgtaatccgtgggaggggctagaagatgcttcgttt
 gatcaagtggcacgttgggttaaaaagtgtgagggatgttgacatttatcctaagcgtct
 actattcacagagcgttagtaataaagaaatagattttaaagattttgaagaacgggtg
 45 gatgaaaatcgtgcgcattataataataggtcactatctgacagtgatataacacatat
 attcaagagcgaaaaaattttaaaccgattgaagaagggtactttaatacaaaaagataac
 gagaaactggaaaaatgggtacaaactaattttaaggattataagaaaaaagaagatgtg
 atagcgcaaaagtgttagtgtttcacaaaggaaggtacacgacttattgatattttaaat
 gctcatatgattaagtgtgctaaattatatgttgatgactttcaatcaagttggactatg
 50 ccaaaaagagaaaaaggaattctatcatgcctggcaacgtttagttaaacatgatccatta
 ttcacaaaaaaacaacgacttactttagcacatttgccaaatcaagcaaccgaagcaata
 gagtacgcctttcaagaattaggagtaaaaagaagaacatcgacaatcatataattgagagt
 catttattatctttaccaggttgggcaggaatcatgtatcatcggtcacagacacaaagt
 aatgatgcgtacttattaacagactatgttgcgattcgtctatcaattgagatgggtactt
 55 ttaaatgaccaccatacaacattatttaaaaaaatctatagcccagttaaagtga

Sequence 2634

MLQSDINELVNQAKRVITPLSPISTFAARNPWEGLDASFDQVARWLKSVRDVDIYPNAS
 TIHRAISNKEIDLKVFEEERLDENRAHYNNRSLSDSDINTYIQRANKLKTIEEGYFNFKDN

EKLEKWVQTNFKDYKKKEDVIAQSASVFTKEGTRLIDILNAHMIKWSKLYVDDFQSSWTM
PKREKGFYHAWQRLVKHDP LFTKKQRLTLAHLPNQATEAIEYAFQELGVKEEHRQSYIES
HLLSLPGWAGIMYHRSQTQSNDAYLLTDYVAIRLSIEMVLLNDHHTTLLKKSIAQLK*

5 Sequence 2635

Contig_0787_pos_0_876,

is similar to (with p-value 0.0e+00)

>pir:pir|A40585|A40585 DNA topoisomerase (ATP-hydrolyzing) (EC 5.99.1.3) chain B - Staphylococcus aureus >gp:gp|X71437|S
10 AGYRREC_2 S.aureus genes gyrB, gyrA and recF (partial). NID: g296393. >gp:gp|D10489|STAGYRABA_1 Staphylococcus aureus genes for DNA gyrase A and B, complete cds. NID: g540540.
gtgatggtgaatacattgtcagatgtaaacacacagataattatggtgctggacagata
caagttttagaaggtctcgaagcgggttcgtaaaagaccgggtatgtatattggttcaact
15 tcagaaagaggggttcacatttagtatgggaaattgttgataatagattgacgaggca
ttagcaggttatgctagtcattgaagttgtaattgagaaagacaattggattaaagtt
actgacaatggcgtggtattcctgttgatattcaagaaaagatgggacgccctgctgtc
gaagttatcttaactgtacttcacgctggaggttaattcggaggtggcggatacaaagta
tctggcgggtcttcacggtgttggtatcttcagttgtaattgcactctcacaagatcttgaa
20 gtttatgtacatcgtaatggcagatttatcatcaagcctataaacaaggtgtgccacaa
tttgatcttaaaagaaattggcgatacagataaaacaggtacagctattcgattcaaagcc
gataaagaaatctttacagagacaacagtttataactatgaaacacttcaaaagcgtata
cgtgagcttgctttcttaataaaggtattcaaattactttaaaagatgaaagagaagag
gaagttagagaagactcatatcattatgaaggcgggattaaatcctatgtagatttatta
25 aatgagaataaagaacctcttcacgatgaacctatataatccatcagtcctaaagcagat
attgaagtggaaattgcacttcaatataacagtggtatgcaaccaacttattaacgtat
gcgaataatattcatacatacgaggggtggtatTCAA

Sequence 2636

30 VMVNTLSDVNNNTNYGAGQIQVLEGLEAVRKRPGMYIGSTSERGLHHLVWEIVDNSIDEA
LAGYASHIEVIEKDNWIKVTDNGRGIPVDIQEKMGRPAVEVILTTLHAGGKFGGGGYKV
SGGLSGVSSVNVNALSQDLEVYVHRNGTIYHQAYKQGVPOFDLKEIGTDKTGTATRFKA
DKEIFTETTVYNYETLQKRIRELAFNLKGIQITLKDEREEVREDSYHYEGGIKSIVDLL
NENKEPLHDEPIYIHQSKDDIEVEIALQYNSGYATNLLTYANNIHTYEGGIQ

35

Sequence 2637

Contig_0788_pos_5197_6411,

is similar to (with p-value 3.0e-17)

>pir:pir|I64093|I64093 ribosomal protein S14 (rpS14) homolog
40 - Haemophilus influenzae (strain Rd KW20) >gp:gp|U32762|U32
762_5 Haemophilus influenzae Rd section 77 of 163 of the complete genome. NID: g1573797.
gtggatgatgtgacaaaatattggtccagttgatggagatccgatcacgtcaacggaagaa
attccattcgacaagaaacgtgaattcaatcctgatttaaaaccaggtgaagagcgtgtt
45 aaacaaaaaggtgaaccaggaacaaaaacaattacaacaccaacaactaagaaccatta
acaggggaaaaagttggcgaaggtgaaccaacagaaaaataacaaaaacaaccagtagat
gaaatcacagaatatggtggcgaagaaatcaagccaggccataaggatgaatttgatcca
aatgcaccgaaaggttagccaagaggacgttccaggtaaaccaggagttaaaaaccctgat
acaggcgaagtagtcacaccaccagtggtgatgtgacaaaatattggtccagttgatgga
50 gatccgatcacgtcaacggaagaaattccattcgacaagaaacgtgaattcaatcctgat
tttaaaaccaggtaaagagcgcgttaaacagaaaggtgaaccaggaacaaaaacaattaca
acaccaaacaactaagaaccatttaacaggggaaaaagttggcgaaggtgaaccaaagaa
aaagtaacaaaaacaaccagtagatgaaatcacagaatatggtggcgaagaaatcaagcca
ggccataaggatgaatttgatccaaatgcaccgaaaggttagccaagaggacgttccaggt
55 aaaccaggagttaaaaatcctgatacagggcgaagtagttactccaccagtggtgatgtg
acaaaatattggtccagttgatggagatccgattacgtcaacggaagaaattccgtttgat
aaaaaacgcgaatttgatccaaacttagcgccagggtacagagaaagtcgttcaaaaaggt
gaaccaggaacaaaaacaattacaacaccaacaactaagaaccctatggcgaagaaatct
aaaatagcaaaagaacaaaaagacaagaattagtaataaatattacgagttacgtaaa

gaattaaaagcaaaaggggactatgaagcattaagaaagttgccaagagattcatctcca
actagattaactagaagatgtaaagtaactgtagacctagaggtgtgttacgtaaat
gaaatgtctagaattgcatttagagaacatgcgcataaaggtcaaattccagggtgtaaaa
aaatctagttggttaa

5

Sequence 2638

VDDVTKYGPVDGDPITSTEEIPFDKKREFNPDLKPGEERVQKQKGEPGTKTITPTTKNPL
TGEKVGEGETEKITKQPVDEITEYGGEEIKPGHKDEFDPNAPKGSQEDVPGKPGVKNPD
TGEVVTTPVDDVTKYGPVDGDPITSTEEIPFDKKREFNPDLKPGKERVQKQKGEPGTKTIT
10 TPTTKNPLTGEKVGEGETEKITKQPVDEITEYGGEEIKPGHKDEFDPNAPKGSQEDVPG
KPGVKNPDTGEVVTTPVDDVTKYGPVDGDPITSTEEIPFDKKREFDPNLAPGTEKVVQKG
EPGTKTITPTTKNPMKKSIAKEQKRQELVNKYELRKELKAKGDYEALRKLPDSSP
TRLTRRCKVTGRPRGVLRKFEMSRIAFREHAHKGQIPGVKKSSW*

15 Sequence 2639

Contig_0788_pos_5829_4909,

putative peptide of unknown function

gtgatttcatctactggttgtttgttactttttctgttggttcaccttcgccaactttt
tccctgttaataaggttcttagttgttgggtgttgaattgttttctgttccctggttcacct
20 ttctgttttaacgcgctctttacctggttttaaatcaggattgaattcacgtttctgtcg
aatggaatttcttccgttgacgtgatcggatctccatcaactggaccatattttgtcaca
tcatccactggtggtgtgactacttcgctgtatcagggttttaactcctggtttacct
ggaacgtcctcttggctacctttcgggtgcatgttgatcaaattcatccttatggcctggc
25 ttgatctcttcgccaccatattctgtgatttcatctactggttgtttgttatttcttct
gttgggttcaccttcgccaactttttccctgttaatgggttcttagttgttgggtgttga
attgtttttgttcttgggttcacctttttgtttaacacgctcttcacctggttttaaatca
ggattgaattcacgtttcttgtcgaatggaatttcttccgttgacgtgatcggatctcca
tcaactggaccatattttgtcacatcatccacagggtggagtaactacttcgctgtatca
ggatttttaacccccggcttacctggttgcgttgtttgactacctttcgggtgcatgttg
30 atcaaattcatccttatggcctggcttgatttcttcgccaccataatgaacgatttcac
cactggttgtttgttattttttctgttgggttcaccttcgccaactttttctcctgtatt
aggattgacataagttggtgttgttgttgtttcaattcctggttcacctttttggactac
ttttctgtacctggggctaa

35 Sequence 2640

VISSTGCFVTFVSGSPSTFSPVNGFLVVGVVIVFVPGSPFCLTRSLPGFKSGLNSRFLS
NGISSVDVIGSPSTGPYFVTSSSTGGVTTSPVSGFLTPLPGTSSWLPFGAFGSNSSLWPG
LISSPPYSVISSTGCFVIFSVGSPSPSTFSPVNGFLVVGVVIVFVPGSPFCLTRSSPGFKS
GLNSRFLSNGISSVDVIGSPSTGPYFVTSSSTGGVTTSPVSGFLTPLRLTLWLRLCLTFRCIW
40 IKFILMAWLDFFATIMNDFIHWLFCYFCWFTEFANFFSCIRIDISWCCCFNSWFTFLDY
FFCTWG*

Sequence 2641

Contig_0788_pos_3383_2700,

45 is similar to (with p-value 2.0e-16)

>sp:P54176|HLY3_BACCE HEMOLYSIN III (HLY-III). >pir:pir|S
59967|S59967 hemolysin III - Bacillus cereus >pir:pir|S52296
|S52296 hemolytic factor - Bacillus cereus >gp:gp|X84058|BCH
MLYSN_1 B.cereus gene for hemolytic factor. NID: g662879.

50 atgtctcaatcatctaagcgaaaaacgattctgtgtagaaacttttaaggacatcatt
cctttaacatttggagaagagatttggaatgcagcatcacatggtgctgctgcattactt
acattatttattcttccatattgcagcagtgcatagttttaacaatggtggcacattagag
tcaattagtgtgcagtttatgtgattagtttttatgatgttcatatcatcaaccatt
tatcattcaatgcaaaataatacgtctcatataatataattaaggattattgaccatagt
55 atgatttatgtggctattttcaggaaacatacacacctgtttgttaagtgttgcggcggt
tggttaggttggtcgtgacaatattattatgggggaacgacattatgggggattttgtat
aaatcaatagcaactaaagtcaatcatagattaagtctcattgtttattttggtgatggga
tggttaggtatcatatttttacctattattattatgcgaacatcatggtggtttattttc
tttatatttcttgggtgggttatcttatactatcggagcatggttttatgccccaaaaaat

aaaccttattttcatatgatttggcatatatttattgttcttgcttctttcttacatg
ataggcattttttattttatgtga

Sequence 2642

- 5 MSQSSKRKNDSSVETFKDIIPLTFGEEIGNAASHGAAALLTLFILPYAAVHSFNNGGTLE
SISVSVYVISIFMMFISSTIYHSMQNNTSHKYILRIIDHSMIYVAISGTYTPVLLSVVGG
WLGWLVTILLWGTTTLWGILYKSIATKVNHRSLIVYLVMGWVGIIIFLPIIIMRTSWWFIF
FIFLGLS YTIGAWFYAQKNKPYFHMIWHIFIVLASFLHMIGIFYFM*

10 Sequence 2643

Contig_0788_pos_2461_1124,

is similar to (with p-value 7.0e-20)

>gp:gp|D50098|D50098_1 Bacillus subtilis macronuclear DNA fo
r multidrug transporter, complete cds. NID: g1856976.

- 15 atgaatcttaagtctatcattactgtaatggcactcataactaataatgtttatggagct
atagaaacatctattatttcatttagcattaccaacaataaaaaatagtttgaatgccggt
aatctagtttcatttagtatttaccgtatattttattgcttttagtcatagctaaccctatc
gttggtgaacttatgtctagattttaaattatttaccattgctgtttaggggtattattg
tttgccttaggttagtttaattgtcgggatttaagtcagacgtttacttttttaattatctct
20 cgaacagtacaagggttttggagcaggagttatgatgtcactctcacaaatagttcctaag
ttggcttttgaaattcctttgagatataaaattatgggtatagttggaagtgtttgggga
atttcgagtattatttggccattattaggtggtgcgatttttagagtttgcttcattggcat
tggtattctatatcaatattcctattgctatagtggaataataacttgtagcttatgact
tttcattttcctgatgagacacaagtcacaacagagtcgttttgatataaaaggattgatt
25 atcttttatattctttatagcttttattaatgtttggtttactcaaccaacatcatattatt
tttaacatattctcaatttttttagcttttagctgttttatggctactattttaaatagaa
aatagtatcgaacaaccattttctccaacaaaagaatttaacatatcaatagttcctagtt
tttataacggattttacttattgagataacactgatgggatataatttatataaccagta
tatttacaagaaaaacttagtttatcacctttacaaagtggatttgtaatttcccggttg
30 tctgttgcttggtattacgcttaatttcaatttagctaaaatagaagcgcattttactaga
aaaacattatatatttgctcattttttgttttattagtttagtagtctgatgataatgttt
ggcttaaaactcccattgcttattgcttttgcggttggttttgcagggttaagttttggt
tatatttatatacaaaaagatagtggttattgtccaagaggaaacttctccaaaaaatatgaaa
aagatgatgtcattttatgcattgacaaaaaatttaggttcgtcagtcggatctacgatt
35 atgggctatatgtatgcactaaatgttggtttatttgggttctaatttacacaatgatta
ggattagtccttaataattgcagtatgtttaattgtaattgtggatgacattatataaaagc
aatactattcaatcttag

Sequence 2644

- 40 MNLKSIITVMALILIMFMAAIETSIISLALPTIKNSLNAGNLVSLVFTVYFIALVIANPI
VGELMSRFKIIYIAVVGVLFLALGSLMSGLSQTFTFLIISRTVQGFAGVMMSSLSQIVPK
LAFEIPLRYKIMGIVGSVWGISSIIIGPLLGAILEFASWHWLFYINIPAIIVAILVLMT
FHFPEDETQVQSRFDIKGLIIFYIFIALLMFGLLNQHHIIFNIFSIILALAVLWLLFKIE
NSIEQPFPLPTEKNISIVLVFITDLLIAITLMGYNLYIPVYLQEKLSLSPLOSQGFVIFPL
45 SVAWITLNFNLAKIEAHFTRKTLIYCSFFVLLVSSLMIMFGLKLPLLIAFAVVFAGLSFG
YIYTKDSVIVQEETSPKNMKMMSFYALTKNLGSSVSGSTIMGYMYALNVGLFGSNLHNVL
GLVLIIVAVCLIVMWMPLYKSNTIQS*

Sequence 2645

- 50 Contig_0795_pos_9882_9001,

putative peptide of unknown function

- atgccaaagtccaattgttgcattgagcatggatggagcgggtatcaatttcaattcaccattt
gcagatattggcaattttatttaggtcttaactggttagcaatattacttttatatggaagca
gtt.gtgtaaccgttttggtactggagtttcttttgggttactggacgtgtgttacgc
55 gctatggaagaaaatgggcatattcctaaattcttaggtaaaattaataaaaagtataat
atcccacgtgttgccattgcatttaattgcaattatcagcatgattatggtgacattgttc
cgtgactgggttacactagctgcggttatttctactgcaacattagttgcatatttaact
ggtccaactacggttatttcattacgtaaaatggcaccaaaaatgactcgtccattttaa
gctaataattttaaaatttatggcacctttatcctttgttttagcatcattagctatctat

tgggcaatgtggccaacaacagcagaagttattttaattatttttaggtttacctatt
 tttttcttctatgaatataaaatgaattggaagaataactaaaaacaaattggcggaagc
 ttatggattattatctatcttattgttctcgcattcttatcattttattggaagtaaagag
 5 ttcaaagccttaaattggattcactatccatgggatttcttagtcattgtaatcgttgct
 ttaatcttctatcaactaggtacgacaagttactttgaaagtattttatttcaagcgtgcg
 aacaagttgaataagaaaatggcgataagttgcgtaaaacacgcaaaaaagcgcgatcat
 aaagatrggaagaacgcgatcgacaagagcaaaatcaataa

Sequence 2646

10 MPSSMLHEHGWSGINFNSPFADMAILLGLNLAILLYMEAVVSPFGTGVSVFVAVTGRVLR
 AMEENGHHPKFLGKINKKYNIIPRAIAFNAIISMIMVTLFRDWGTLAAVISTATLVAYLT
 GPTTVISLRKMAPKMTTRPFKANILKFMAPLSFVLASLAIYWAMWPTTAEVILIIILGLPI
 YFFYEYKMNWKNKTKKQIGGSLWIIYLIIVLAFLSFIGSKEFKGLNWIHYPDWFLVIVIVA
 15 LIFYQLGTTTSYFESIYFKRANKLNKKMGDKLRKTRKKARHKDWKERDRQEONQ*

Sequence 2647

Contig_0795_pos_8696_7329,
 is similar to (with p-value 2.0e-47)
 >gp:gp|U46134|BSU46134_3 Bacillus subtilis putative orf1 unk
 20 nown protein, putative transcriptional regulator (slr), and
 intracellular esterase B (estB) genes, complete cds. NID: g1
 762123. >gp:gp|Z99121|BSUB0018_125 Bacillus subtilis complet
 e genome (section 18 of 21): from 3399551 to 3609060. NID: g
 2635827. >gp:gp|Z71928|BSYVEFGNS_1 B.subtilis pnbA, sigL, yv
 25 e[J,K,L,M,N,O,P,Q,R,S,T] and yvf[A,B,C,D,E,F,G,H] genes. NID
 : g1495276. >gp:gp|Z94043|BSZ94043_47 B.subtilis genomic DNA
 fragment (88 kb). NID: g1945641.
 gtgtgttcaaacatggttcaagtcaagataggttaactgtaccatcaatggcttacacaaa
 aagaatattgatgtattcttaggtatcccgatgccaaaccattcaataagatatctcga
 30 ttccaacattcaaagtttatggaactaagcaaaccaatgattgatgcaactcatattcaa
 tccatcccaccacaaccctacaattcacttgaagacttttttcgatgacagattcatcg
 ttaattcttttaacaaaaatgattattgcctgttttaaatatttggaaccatcgccc
 aatcaaaatcatttacctgtagtgtttacttttatgggtgtagttttcttcaaggacat
 ggcacagctgaactatattgtcctgaacacacatagtagaacaagaaaatataatagtagtt
 35 acttttaattatcgcttaggtgcactcggctacctagattggtcttattttaatcaacat
 ttgaactataataatgggattttctgatcaaattaacgttttaagatgggtacatcaatat
 atcgaacattttggcggtgattcaaataacgtgacactaatgggtcaatctgcaggtagt
 atgagtatcatgacattaatgcaaatgcccgaacttgatgattattatcataaagtgatg
 ttattaagtgaacgttaactactgatacaccactcaatgcacatactaaagtacaacat
 40 ttttcacaactcatgaggcattttttcctaataaaacacttaagacacttaccagtgat
 gacattttatatctaatggagctctcaaaaaatagagcgtggaagatctcgtggacttgat
 ttgatttatcaacctattaaagatcatcatatgtcacgatccattaaaaaatttcccaaa
 ccgacattcatgagttatacacacgatgaaggtgatatttatattgaagacgcaaacgc
 accttaccttctgaacgttttattcacttgatgtctcaatatggtacacacgtcgaaaaa
 45 aatgatgccctcacaatgaacaacaaagaaatttaataacagagtattgtttgttcgt
 ccaatttatctatttttaaatcaaataagttgacgacacttggttagcacgttttgat
 tggcaccacccatacctcctacttttaaaagtgcatacatatattggatttagtattt
 tggtttggtcactctctattttgactaaaaatcattattctataactcaacatgatatg
 aatttaagtagtaacatgatatctgacttagcttattttgcccgaaggttaagatgcc
 50 tggaaatgttatgaacctcaacatcaagcgttacatatctatggataa

Sequence 2648

VCSNMVQVKIGNCTINGLHKKNIDVFLGIPYAKPFNFKISRQHSKFMELSKPMIDATHIQ
 SIPPQPYNSLEDFFSMTDSSFNSFKQNDYCLFLNIWKPSSNQNLHPVVIYFYGGSLQGH
 55 GTAELYCEHIVEQENIIIVTFNYRLGALGYLDWSYFNQHLNNGISDQINVLRVWHQY
 IEHFGGDSNNVTLMGQSAGSMSIMTLMQMPLEDDYHKVMLLSGLTLDTPLNHTKVQH
 FSQLMRHYFPNKTLLTSLDDILYLMESQKIERGRSRLDLIYQPIKDHMSRSIKKFPK
 PTFMSYTHDEGDIYIEDATRTLPSERFIHLMSQYGTHVEKNDALTMKQQRNLITEYCFVR
 PIYFLNLQMNSCDTWLARFDWHQPHTSYFKSAYHILDLVFWFGHLISILTKNHYSITQHDM

NLSSNMISDLAYFARKGKMPWKCYEPQHQLHIYG*

Sequence 2649

Contig_0795_pos_2619_2143,

- 5 is similar to (with p-value 4.0e-68)
 >gp:gp|Z99108|BSUB0005_72 Bacillus subtilis complete genome
 (section 5 of 21): from 802821 to 1011250. NID: g2633055. >g
 p:gp|D78509|D78509_8 Bacillus subtilis YfjG-YfjR genes, complete cds. NID: g2780390.
- 10 atggaagacgtgacagatattgtctttcggcatgttgtcagtgaaagctgcgagaccagat
 gtattttttactgaatttaccataactgagagttactgtcaccctgaaggtattcatagt
 gtgcgcggacgcttaacttttagtgacgacgaacaaccaatggtagcgacacatctggggc
 gataaaccagaacaattccgagaaatgagtatcggttagcgagatgggttttaaaggt
 atagattttaaataatgggttgcctgtcgcaaacgttgcgaaaaaggtaaaggatccggc
 15 ttaattctacgacctgaaacggcagccgaaatcattcaagcttctaaagcaggtggtcta
 ccggtcagtgtaaaaaacagcttttaggttattacgatatcgatgaatggcgagactggtta
 aaacacgtcttcgaacaagtttaggtgcgcgctctggtttaatggcagagccaaatga

Sequence 2650

- 20 MEDVTDIVFRHVSEAAARPDVFFTEFTNTESYCHPEGIHSVRGRLTFSDDDEQPMVAHIWG
 DKPEQFREMSIGLADMGFKGIDLNMGCPVANVAKKKGSGGLILRPETA AEIIQASKAGGL
 PVSVKTRLGYYDIDEWRDLKHVFEQVRCALWFENGRAK*

Sequence 2651

- 25 Contig_0795_pos_2063_660,
 is similar to (with p-value 0.0e+00)
 >gp:gp|AF029225|AF029225_1 Staphylococcus carnosus NarG, Nar
 H, NarJ, and NarI genes, complete cds. NID: g3929521.
- 30 atgactgcgacgccattatattcagatatcgttttacctgctgcaacttggtatgaaaaa
 catgatttatcttctacagacatgcatccatttattcatccatttaaccagcgattgac
 ccattatgggaatcgcttcggactgggatatttataaaaactctaagtaaaagctgtttca
 gaaatggcgaaagattatcttccaggtaaatttaaagatgtcgttaactacaccattagga
 catgattcaaaacaagaaatttcaactgaatacgggtattgtaaaagattggtctaaagga
 gaaattgaagggtgtgccaaggtaaaacaatgcctaatttttctatcgtagagcgagactat
 35 acacaaatttacgataaaattcggttactgttggtccaaaactagaaaaagggaaaaatagg
 gctcatggtgtgagttatagcggttagtgaagagtacgaagaacttaaaagtatagttgga
 acttcgaatgatgataataactatttcagttaaaaatgatagaccgagaatagatacagcg
 agaaaaactagcagatgtcattttgaatatatcctctgctacaaaacggcaaatattacaaa
 aagtcatatgaagatttagaaaaatcaaacaggtatggaacttaaaagatatattctaaagaa
 40 cgtgcttctgaaaagatatcattcttaaacattacttctcaaccaagagaagtgattcca
 actgcagtatctccctggtcttaataaagatggaagacgctactcaccgtttacaactaat
 gttgaacggttagtgccatttagaacactaaactggacgtcaaagttattatatagatcat
 gaggtattccaacagtttggcgaaagtttaccggtatataaacctactttacctccaatg
 gtatttggtgctcgtgataaaaaagttaaaggtggacaagatacattagtgtctcgatac
 45 cttacacctcatggaaaaatggaatattcattcaacttatcaagataatgaacgcagtggtg
 acgttggttagagggtggaccagttgtatggatttcaaatgaagacgcagctgacctgggt
 attaatgataacgactggttagaagtatacaacagaaacggagttgttactgccagagct
 gtaacatctcatcgtagcctagaggcacaatgtttatgtatcatgcacaagataaacat
 atagagacacctggttctgaaattactgatactcgtaggaggttctcataatgcacctact
 50 cgtattcacttgaaacctactcaattagtaggaggatagcacaattagttatcacttt
 aactattatggaccaatttgaaatcaaagagatgagtatgtagctgttagaaaaatgaag
 gaggtcaattggcttgaagattaa

Sequence 2652

- 55 MTATPLYSDIVLPAATWYEKHDLSSTDMHPFIHPFNPAIDPLWESRSDWDIYKTLKAVS
 EMADYLPKFKDVTTPGLHDSKQEI STEYGIVKDWSKGEIEGVPGKTMNFSIVERDY
 TQ:YDKFTVGPKEKKGIGAHGVSYSVSEYEELKSIVGTWDDNTISVKNDRPI:IDTA
 RKVADVILNISSATNGKLSQKSYEDLENQTMELKDISKERASEKISFLNITSQPREVIP
 TAVFPGSNKDGRRYSPTTNVERLVPFRILTGRQSYIIDHEVFQQFGESLPVYKPTLPPM

VFGARDKKVKGGQDTLVRLYLTPHGKWNHSTYQDNERMLTLFRGGPVVWISNEDAADHG
 INDNDWLEVYNRNGVVTARAVTSHRMPRGTMFMYHAQDKHIETPGSEITDTRGGSHNAPT
 RIHLKPTQLVGGYAQISYHFNYGPIGNQRDEYVAVRKMKEVNWLED*

5 Sequence 2653

Contig_0795_pos_0_646,

is similar to (with p-value 0.0e+00)

>gp:gp|AF029225|AF029225_2 Staphylococcus carnosus NarG, NarH, NarJ, and NarI genes, complete cds. NID: g3929521.

10 atgglaattgaatctagacaaatgtattggtgtcacttgcagtgtgacatgtaaaac
 acatggacaaatcgacctggtgcagaatataatgtggtttaataacgtagaaacaaaaccg
 ggtgtaggatatccaaaaagatgggaagaccaaggacaataataaagggtggttggtgcta
 aataaaaaaggaaagccttgaattaaaatctggaacagatggtaaaaaattgctttagg
 15 aaaatcttctataatccagacatgccactcattcaagattattatgaaccgtggacatat
 aactatgaacacttaaccaatgctaaacaaggacagcactctcccgtagacagctcac
 tctttaatttcaggtgatagattgaatcttaaatgggggccaaactgggaagatgattta
 gctggaggtcacattacaggaccagaggatccaaatattcagaaaatagaagaagatatt
 aaattccaattcgatgagacatttatgatgtatttaccaagactatgtgaacactgttta
 aatccaagttgcgtagcatctgtccatcaggagctatgtataaacgagatgaggatggt
 20 atcgtaactcgatcaagaagcctgtcgaggttgagataactgta

Sequence 2654

MVLNLDKICGHTCSVTCKNTWTNRPGAEMYWENNVEKPGVGYPKRWEDQGQYKGGWVL
 NKKGKLELKSGNRWSKIALGKIFYNPDMPLIQDYEPWTYNYEHLTNAKQGQHSPVATAH
 25 SLISGDRNLKWPWEDDLAGGHITGPEDPNIQKTEEDIKFQFDETFMMLPRLCEHCL
 NPSCVASCPSGAMYKRDEDGIVLVDQEACRGWRYCX

Sequence 2655

Contig_0796_pos_1625_2248,

putative peptide of unknown function

30 atggattatgtatacacaaatataaaaaatcctagatataacatcattcaaaaagataat
 cgctatttaattggtcgatttagaacaataattggtattcatatttatgtccaatgctgaat
 tggtttatacctattaaattcacagaatttaacttatcaagaattcaataatataaacata
 tttcataatggaggacaaaagagtcattggtatgatggctgctggcgttggcgtcactatc
 35 agtgtgctattaagaagctctgtgggttatatagatattaatattagtcgaatttgata
 gtttttatgtttttaattggatttgttgctgtgatcacacttcgtttatctataagaaag
 aagttaaatcatccagcatttaataaaaaagagtaaaacaaaagtaaatattgataccatca
 tttaaaaatgatattggtggtgttttgcattttatgatgctgtttttctcaattgca
 ccttttcaaattgattttgaggaaaagaaaacatcttaggatataactttgggtagggt
 40 gtattatttatattactactttgaatatggcttcaatttctgatagaaaagtacatgcc
 aaaattaaaaatataagaagatag

Sequence 2656

MDYVYTIYKNPRYNIIQKDNRYLMVDLEQNWYSYLCPLNWFIPKFTELTYQEFNNINI
 45 FHNGGQKSHGMMAGVGVTISVLLRSLVGYIDINISRIWIVFMFLIGFVAVITLRLSIRK
 KLNHPAFENKSKQKVILIPSFKNMILVFCYFMMFLFFSIAPFQMI FEEKKNILGYILWVG
 VLFIFTTLNMAISIDRKVHAKIKNIRR*

Sequence 2657

Contig_0796_pos_4131_4508,

putative peptide of unknown function

50 atgcaattttattatagtaattgtgaacagaatgcataaaacttattagtcacatcaacgta
 caacctaatgaaggattttctttatgtgtgaatggtaagaaaagtaatacaaaataatgaa
 atgcaaaaagtgaaagctttcttatactatgccgattaaagataaaaatgaacacagttgat
 55 gcatatgaaaatcttatttacgatacatatttggaagaacaaacaaaatttacgcattgg
 gaagaattaaaaattcttgaaattttattgatgatattgaaaatgtatggaaacaagaat
 agccacagtttcttaattatgcctttggatgctatgggcctaaagaaagtgaataattac
 ttatgtgaagacggattga

Sequence 2658

MQFYYSNCEQNASNLLVINVPNEGFSLCVNGKKSQNQNMQKVLSYTMPKDKMNTVD
 AYENLIYDTLIGEQTFTHWEELKILGNLLMILKMYGNKNSHSFLIMPLDAMGLKKVKNY
 LVKTD*

5

Sequence 2659

Contig_0796_pos_4627_0,

is similar to (with p-value 6.0e-53)

>gp:gp|L76359|STMDRRC_1 Streptomyces peucetius daunorubicin
 resistance protein (drrC) gene, complete cds. NID: g1196906.
 atggatttttattaatattacaggtgcttcacaaaataacttgaaaaacatagatgtaaat
 atcccaaaacacttagtaacggtattttacaggtcgcttctggttcagggaaatcatcttta
 gtgtttaatactgttgctgcggagctcgaacagctactaaatgaaagttattctagttaa
 attcaattttcatttaaatcaacaacccagaccgaaagtaaaagaaaattaaaaatcttct
 gtagcaatgacgattaatcagaaaagattcaatgggaattctcgctccacggttaggaaca
 gtttcagatatatatgcttctgtagattactgtggtctagaataggcgaaccgtttgtt
 ggttattcagatgcatattccttcaatagtcctaagggcattgtgtaaaacttgtgagggg
 ttaggatataattgaagacattaaacttagatgaattgctagattgggataagtctttaaat
 gaaggtgcaatagactttcttcttttgaccagacaaagagcgtggtaaagcctatcga
 gatagt

20

Sequence 2660

MDFINITGASQNNLKNIDVNI PKHLVTVFTGRSGSGKSSLVFNTVAAESEQLLNESSYSSY
 IQFHLNQPRPKVKKIKNLVAMTINQKRFGNSRSTVGTVSDIYASVRLWSRIGEPFV
 GYSDAYSFNSPKGMCKTCEGLGYIEDINLDELDDWDKSLNEGAIDFPSFGPKERKAYR
 DS

25

Sequence 2661

Contig_0796_pos_728_342,

putative peptide of unknown function

atggctattgtaaataaggttattattgttgaaaggttaaatacggataagaaaagagtacaa
 caagtaatcgctgaacctgcaaatatcatttgtacacatggcactatgagtatagataag
 atagacaacatgatagaaacactttatgacaaacaagtttatgttcttgccgattctgat
 gatgagggtgaaaaaattagaaaaatggtttaaacggtattttaagcgaaagtgaacataat
 tatgttgataaaacgttttgtgaggttgctaagtgctcctaaaaattatttagcacatgta
 ttaagtagatatggttttaattgtaaaaaaagaaaagaaacttatgaataatttaaaaact
 gaaaggctagtttttagtaaatgaataa

35

Sequence 2662

MAVNVKVTIVEGKSDKKRVQVIAEPANI ICTHGTMSIDKIDNMETLYDKQVYVLA DSD
 DEGEKIRKWFKRYLSESEHIYVDKTFCEVAKCPKNYLAHVLSRYGFNVKKEKKLMNNLKT
 ERLVLVNE*

40

Sequence 2663

Contig_0798_pos_5428_4553,

putative peptide of unknown function

atgatattgaattcaaaagttaaaggtattattgctatattgatttcagctgtgggtttt
 agttttatgtcagtccttttttagattggccggtgatttaccagtcctttcaaaaatctcta
 gctagaaaattttgtagccatgtttataaccattattttttatttataaatataggcaacct
 atgtttggaaaattaagtagtcaaccctactcatctcacgttcaacacttgggttaatt
 ggtgtcttacttaatatctacgcaattgatcacatgggtattaagtgtgctgatacatta
 atgaaattaaatcctttttggacaattgttcttagtttaattttttacatgaaaaggta
 cgaaaatatcaaatcacggcgatgattattgctataatagggtgctattaattgttaaa
 ccagaattttcatcatcagttattccttcaatagcaggattactatccgggtatttttgca
 gcttctgcctacacatgtgttagagcactcagcactcgtgaaaaaccttatacgatagtg
 ttttatttttcattattctcagttgtagttcttatacctttttcaatatttacttataca
 cctatgacaacaattcaaatcttttcttactcggcgctggattatcagcagctgtagga
 caaattgggtataacattggcttatagttttgtccagcaaaaagatatctccatcttcaca
 tatgcgtctataatatttactgcattatttggatttattctgtttggagaatcacctgat

45

50

55

atgtttgcaacagtaggatataattgtcattatcgagcaagttactatatgtttgataaa
gcaagacgtgaaacaactataaatcaaaataattaa

Sequence 2664

5 MILNSKVKGIITAILISAVGFSFMSVFFERLAGDLPVFQKSLARNEVAMFIPLFFIYKYRQP
MFGKLSSQPLLISRSTLGLIGVLLNIYAIDHMLSDADTLMKLNPEWTIVLSLIFLHEKV
RKYQITAMIIATIGMLLIVKPEFSSSVIPSIAGLLSGIFAASAYTCVRALSTREKPYTIV
FYFSLFSVVVLIPFSIFTYTPMTTIQILFLLGAGLSAAVGQIGITLAYSFAPAKDISIFT
YASIIFTALFGFILFGESPDMEFATVGIVIVIIGASYMFDKARRETTINQNN*

10

Sequence 2665

Contig_0798_pos_3629_2250,

is similar to (with p-value 0.0e+00)

15 >sp:sp|P94408|YCLF_BACSU HYPOTHETICAL 53.3 KD PROTEIN IN SFP
-GERKA INTERGENIC REGION. >gp:gp|Z99106|BSUB0003_15 Bacillus
subtilis complete genome (section 3 of 21): from 402751 to
611850. NID: g2632653. >gp:gp|D50453|D50453_69 Bacillus subt
ilis DNA for 25-36 degree region containing the amyE-srfA re
gion, complete cds. NID: g1805369.

20 atgagagcaatgcttatttttatatgtatttgcactccaggaaaacggacttggaaatg
gacaaaacaacagccatgtctatcatgtctgtatatggatcattaatttatatgtcttca
attccaggtggctggatagctgatagaattacaggtacccgaggtgccacattaatcgga
gctatattaattattataggtcacatatgtttaagccttcctttgcaatggtaggttta
ttcacttctatgttctttattattgtaggatcaggtttaatgaaaccaaactttcaaat
25 attgttggtagactctatccagaaaacgacgtacgcgatggatgctggatttgttattctt
tatatgtcagtgaacatgggtgcactcgtttcaccaattattttacaacactatattgat
attagaaatttccatggcggtctcttgattgcagcaatagggatggctcttggtattagt
tggtacttactatttaaccgaaaaactttgggtagtatcggtatgaaaccgacaaaccca
ttatcttctctgaaaagaaaaagtagcgaacaatcatcggaatcggtgttatagcaata
30 gtattaatccttatgattgcttactttacgcatacgctatcatttaatttaacagtaaat
actgttttaatttttaggtattgctttaccaatcatttactttacaacaatgattagaagt
aaaggaagtactgatactgaaagatcaagagtaaaagcatttattccgttattcaattta
ggcatgttattttgggtcaattcaagaacaaggatctaattgtattaaatatctatggatt
gaaaactctgatatgaaattaaattttatttgggtggaaaacacattttgggtgaagctatt
35 ttccaaaccatttaaccattattttattttattttgcacccgtgggtactcttttatgg
caaaagctaggaaagaaacaacctagcctacctaagtttgcaattgggtactatttta
gcaggcgcatcctacatacttatgggagcaatcggtcatatttatggggatacacaattc
tcagtttaactgggttattctttcatatcggtatctgtgttattgggtgagctttgtctctt
ccaactggtagtagtgacagcagttaaattagcacctaaggcatttaacgcacaaatgatg
40 agcctttgggtatttaactaacgcttcagctcaggccatttaacggtacattagttaaatta
attaaaccaatttggtcaaaccaattactttatcttcttaggtgtgttgcaaccgtgatt
acgttaatttatattagcgtttatttccctaagatttctaaagcaatgaaaggtattcggttaa

45 Sequence 2666

MRAMLIFYMYFALQENGLGMDKTTAMSIMSVYGS LIYMSSIPGGWIADRITGTRGATLIG
AILIIIGHICLSLPFAMVGLFTSMFFIIVGSGLMKPNISNIVGRLYPENDVRMDAGFVIF
YMSVNMGALVSPILQHYIDIRNFHGGFLIAAIGMALGLVWYLLFNKRKTLGSIGMKPTNP
LSSSEKKKYGTIIGIVVIAIVLILMIAYFHTLSFNLISNTVLILGIALPIIYFTTMIRS
50 KEVTDTERSrvKAFIPLFILGMLFWSIQEQGSNVLNIYGIENSMDKLNLFGWKTHGEAI
FQTINPLFILLFAPVVTLLWQKLGKKQPSLPKFAIGTILAGASYILMGAIGHIYCQTQF
SVNWVILSYVICVIGELCLSP TGSSAAVKLAPKAFNAQMMSLWLLTNASAQAINGFLVKL
IKPLGQTNFYIFLGVVATVITLILAFIPKISKAMKGIR*

55 Sequence 2667

Contig_0798_pos_1595_672,

is similar to (with p-value 7.0e-45)

>sp:sp|P39074|BMRU_BACSU BMRU PROTEIN. >gp:gp|L25604|BACBMRU
RBE_1 Bacillus subtilis bmrU, multidrug efflux transporter (

bmr) and its regulator (bmrR) genes, complete cds, and branched-chain 2-oxo acid dehydrogenase (bfmB) gene, 3' end. NID: g2558636. >gp|D84432|BACJH642_251 Bacillus subtilis DNA, 283 Kb region containing skin element. NID: g2627063. >gp:gp|Z99116|BSUB0013_111 Bacillus subtilis complete genome (section 13 of 21): from 2395261 to 2613730. NID: g2634723.

5 atgaaacaaccgtataaccatggtgttcttttctatcatgaacatagcggtttgaaagat
 atacataatggcatagaggagaagttgcaaaatctcttagttcaatgtgtaaacacctctct
 cttcaactcagtgaaaataaaggcgatattattaaatattgtaaatctattaaaaatgaa
 10 aattatagctctgatgtagacggttttatttttaggtggagatggtacacttaatgaa
 ctagttaaattggcggttagcagtatcagttaaatttaccatcggtgtaataccaggtggt
 acctttaacgattttacaaaaacacttcaactgcaccctaattttaaaacagctagtgag
 caattattaacatcacatgctgaatcatatgatgtgttaaaagtgaacgacttatatgta
 ctttaatttcggttgacttggttaatagtacaaaatgcagagaatgttcaagatggttct
 15 aaagatatattcggttaaattcagctatattggatcaaccgttaaaacgttattaaatcct
 gttaaatttgattttctcattgactgttgatggtgaaacaaaagaaggcaataacttcgatg
 atgttaatagcaaacggtcccaatataggtggtggacaaattccgctaaccgatttatcg
 ccacaagatggaagagcaaacacatttgtatttaatatgatcaaacactaaatatattgaat
 gatataattaaaaaacgtgatagtatgaattggaacgaaatcacacaaggtattgatcac
 20 atacaggtatgacatcacactctcaacaaaccctagtatgaaagtggatattgatggc
 gaaatttaatttagaaaacaccaattgagattcaagtattacccaaagcgatacaacttctt
 actgcaactgaacaaaataattaa

Sequence 2668

25 MKQPYNHGVLFYHEHSGLKDIHNGIGEVAKSLSSMCKHLSLQISENKGDI IKYCKSIKNE
 NYSSDQDVLFILGGDGLNELVNGVMQYQLNLP IGVIPGGTFNDFTKTLQLHPNFKTASE
 QLLTSHAESYDVLKVNLDLYVLNFGVLGLIVQNAENVQDGSKDIFGKFSYIGSTVKLLNP
 VKFDFSLTVDGGETKEGNTSMMLIANGPNIGGGQIPLTDLSPQDGRANTFVFNQTLNINL
 DILKKRDSMNWNEITQIDHISGKHITLSTNPSMKVDIDGEINLETPIEIQVLPKAIQLL
 30 TATEQNN*

Sequence 2669

Contig_0798_pos_0_315,
 putative peptide of unknown function

35 atgtttgtgaattatttcacaatatctaaggagtggttgatatgttatctgtaactaaa
 aaaaatacatatgaatcaacaaagatgaagtcacacaaatgattgattcattagcagaa
 aaaggacaagaagctctaaaagaactatctaaaaaatcacacatgagattaatgacatt
 gtacatcagatgagcatggctgctgttgatcagcatatgcatttagctaaactagcttac
 gacgaaacaggtagaggtatttatgaagacaaagctatcaaaaatttatatgcctcagag
 40 tacatatggaattcA

Sequence 2670

MFVNYFTJSKEWLYMLSVTKKNTYESNKDEVTQMIDSLAEKGQEALKELSKKSQHEINDI
 45 VHQMMAAVDQHMHLAKLAYDETGRGIYEDKAIKNLYASEYIWN

Sequence 2671

Contig_0799_pos_5458_5027,
 putative peptide of unknown function

50 atgttcaaaaatatattattaccctatgatttcgaaaatgatttttagtgctatccctgac
 tatttagaaaaagtcaccgatgaagattcagttgttgtaatttatcacggtgtaacagaa
 aatgatcttgcaattagtgtcaagtattataataagcataaaagaagatattattagagaa
 aaagagaaaaaactcactccatttttacgtgaattagaaaaaagagatattcaatataaa
 atagatgtagattttgggcatattaaagatacaatcttagaaaaaattacttctggagat
 ataaaaaatggtgattttgatttagtaattatgagtaaatcatagagtcgatttgatatt
 55 aacatggttttaggagatgttacacataaagattgctaaaaagaagttctgtcccagtacta
 attgttaataaa

Sequence 2672

MFKNILLPYDFENDFSAIPDYLEKVTDDESVVVIYHVVTENDLAISVKYYNKHKEDIIRE

KEKKLT PFLRELEKRD IQYKIDVDFGHIKDTILEKITSGDINNGEFDLVIMSNHRVDLNI
KHVLGDVTHKIAKRSSVPVLIVK*

Sequence 2673

5 Contig_0799_pos_2591_1617,
putative peptide of unknown function
atggaacgattttgtgtgtataatcaaattaactatattcaaataatccggttagaagcc
aaattttaaagcagcgctcctaagatcatggaaaactgatcaggcagatgctcataagctt
10 gcttgttttaggaccgacgctcaaacaacaggcagcttacctatacatgagttaatattc
tttgaattaagagaacgtgcccgttttcatctagaaatcgagaatgaacaaaaatcgactt
aaattttctttatgtttttctatttttactgagagaaataaagtgggtgacacaagattta
tatgactctatttctagagcatattattgtctcgttgatacacaagctaatacaaatatg
attgaacactacgcaggattgaacatgaatgatattaatcttttaagggtaacgcctttt
gatgcgaagtcattaccctaaccaaagtagtcaattgtatgacacttatattggattatgg
15 atagatgtgtttggacgagattgaaatacagagattgtaaacagcttatttcaataatatt
caacataaagatggctataagttgaaaattttaactaagagtagagataatcttacggaa
aatcttatagatgaagttgctcatctcaatgatttatatcaccaagagaaaaaggaaata
agtgtatgaattgaagacgtgatacagaataaaaaagaaacaatcattgatattgaaaca
gtaccgtttgaagaagatcttgaagcgttatttcaaaattaagagttgtagtagattta
20 tctttagagccgaaactttttttacaaatctgttgatttggcgcgggtataccacaaatt
aataaaaaagagaacagattatgttaaacatatgcataatggatatattattgatgacata
tcgcaaacgttagaatcttttagattattttttggcacatttaaaaaatggaattattctt
atgcataattccatga

25 Sequence 2674

MERFCCVNQINYIQMNPLEAKFKTSALRSWKTDQADAHKLACLGPTLKQTGSLPIHELIF
FELRERARFHFLEIENEQNRLKFSLCFSIFTERNKVVTQDLYDSISRAYYCLVDTQANQNM
IEHYAGLNMNDINLLRVTPFDAKSLPNQSSQLYDITYIGLWIDGLDEIEIREIVNSLFQYI
QHKDGYKILKILTKSRDNLTENLIDEVAHLNDLYHQEKKEISDVIEDVIQNKKETIIDIEI
30 VPFEEDLVSVISKLRVVVDSLEPKLFLQICIGAGIPQINKKRTDYVKRMHNGYIIDDI
SQTVESLDYFLAHLKNGIILMHIP*

Sequence 2675

35 Contig_0799_pos_1108_701,
putative peptide of unknown function
atgttcccaccccgaaacacctagtagagatgccactaaccacacctcaacaactcttaca
ttactaggaataccgtcaaataattacccatctcacatacaatttttctgagcatgcatta
ccttggaataagttttatcgctacactatagttttctatcgctattgcaataatctatatt
tatatcgcaaaagaaatatacaaaaatcacactaggttatgggtgctttatttggtagatt
40 atttggattgtttttcatttaattcttaattgccaattatgcatgctgacgaatgctttt
gatcaaccattttcagaacacctatcagaattttttggacacattgtttggatgatggtt
atagaaatgggtcagaaggtattttctataatattcaattaataataataa

Sequence 2676

45 MFPPRTPSRDATNPPQQLQLLGIPSNITHLTYNFSEHALPWISFIVHYSFSIAIAIYI
YIAKKYTKITLGYGALFGIVIVFHLILMPIMHVVPNAFDQPFSEHLEFFGHIVWMMV
IEMVRRYFYNIQLNK*

Sequence 2677

50 Contig_0800_pos_2627_2971,
putative peptide of unknown function
gtgacaaaccggaggaaggtggggatgacgtcaaatacatcatgccccttatgatttgggc
tacacacgtgctacaatggacaatacaaaaggcagcgaaaccgcgaggtcaagcaaatcc
catagaagtgttctcagttcggattgtagtctgcaactcgactatatgaagctggatcg
55 ctagtaactcgtagatcagcatgctacggtgaatacgttcccggtcttgtacacacggcc
cgtcacaccacgagagtttgaacacccgaagccggtggagtaaccatttggagctagcc
gtcgaaggtgggacaaatgattgggggtgaagtcgtaacaaggtag

Sequence 2678

VTNRRKVGMTSNHHAPYDLGYTRATMDNTKGSETARSSKSHKVVLSDDCSLQLDYMKLES
 LVIVDQHATVNTFPLVHTARHTTRVCNTRSRWSNHLELAVEGGTNDWGEVVTR*

Sequence 2679

5 Contig_0802_pos_6032_5214,
 is similar to (with p-value 0.0e+00)
 >sp:sp|P23966|MENB_BACSU NAPHTHOATE SYNTHASE (EC 4.1.3.36) (DIHYDROXYNAPHTHOIC ACID SYNTHETASE) (DHNA SYNTHETASE). >pir:
 pir|A42715|A42715 dihydroxynapthoic acid synthetase - Bacill
 10 us subtilis >gp:gp|M74521|BACMENAQUI_6 Bacillus subtilis men
 aquinone operon, complete cds. NID: g557486. >gp:gp|M74538|B
 ACMENAQOP_4 Bacillus subtilis menaquinone operon: menF, menD
 , menB and menE genes, complete cds. NID: g1185287.
 atgactagacagtgaggaaacacttagagaatatgatgaaattaaatatgaatttttcgaa
 15 gggattgccaaagtaacgattaatcgctccagaagtaagaaatgcatttactcctaaaaca
 gttgctgaaatgattgatgcattttcacgtgcgcgtgatgatcaaatgtatcagtaatt
 gtattaaactgggtgaaggggacaaagcgtttgttcaggtggagatcaaaaaaacgtgga
 cacggtggttatgtagggtgaagatgatattcctcggttaaatgtattagatttacaacgt
 ttaattcggtgtgattcctaaccagtaatatgcaatgggttagaggctatgcaattggtgga
 20 ggaatgtacttaattgtgtttgtgatttaactatcgctgcagacaatgctatttttggga
 caaactggacaaaaagtaggctcatttgatgctgggtacggttctggctacctagctcgt
 atagttggccataaaaaagcaagagaaatctggtacttatgccgtcaatataatgcacag
 gaagctttggatattgggttagtgaaactgtagttccattagaacaagttgaagacgaa
 acagttaaatggtgtaaagacatcatgcaacactcaccaactgctttacgtttcttaaaa
 25 gcagcaatgaatgctgatactgatggttttagctggtttacaacaaatggctggagatgcg
 actttactttactatactactgatgaagcgaaagaaggacgtgacgcgtttaaagaaaa
 cgtaactctgattttgaccaattccctaaattcccataa

Sequence 2680

30 MTRQWETLREYDEIKYEFFEGIAKVTINRPEVRNAFTPKTVAEMIDAFSRARDQNVSVI
 VLTGEGDKAFCSGGDQKQRHGGYVGEDDIPRLNVLDLQRLIRVIPKPVIAMVRGYAIGG
 GNVLVNVCDLTIAADNAIFGQTGPKVGSFDAGYGSGLARIVGHKKAREIWYLCRQYNAQ
 EALDMGLVNTVVPLEQVEDETVMKCKDIMQHSPTALRFLKAAMNADTDGLAGLQQMAGDA
 TLLYYTTDEAKEGRDAFKEKRNPDFDQFPKFP*

Sequence 2681

40 Contig_0802_pos_1667_123,
 is similar to (with p-value 0.0e+00)
 >sp:sp|P39634|ROCA_BACSU 1-PYRROLINE-5-CARBOXYLATE DEHYDROGE
 NASE (EC 1.5.1.12) (P5C DEHYDROGENASE). >pir:pir|S39731|S397
 31 hypothetical protein - Bacillus subtilis >gp:gp|X73124|BS
 GNR_77 B.subtilis genomic region (325 to 333). NID: g413923
 . >gp:gp|Z99123|BSUB0020_74 Bacillus subtilis complete genom
 e (section 20 of 21): from 3798401 to 4010550. NID: g2636240
 45 .
 atggtagtacctttcaaaaatgaacctggtattgatttttcagtacagacaaatgttgag
 cgtttttaaatgaagaattaagggaaagtaaaagcgcaactaggacaagatataccacttg
 attaacggagaaaaacttactaaaactgatacttttaattcagtgaatcctgcgaataca
 tcacagctcattgcgaaagtgtctaaagcaacgcaagatgatattgaaaaagctttcgaa
 50 tcagcaaatcatgcgtatcaatcatggaagaagtggtcgcataaggaccgtgcagaatta
 ctgttacgtgtagccgaattatccgtcgctcgaaaagaggaaatttccgctattatggtt
 tatgaagccggcaagccttgggatgaagcagttggagatgcagctgagggtattgatttt
 atagaatattatgaagatcaatgatggaacttgcagatggtaagccagattatagacaga
 gaaggtgaacataatcgctatttttataaacctattgttacaggcggtgacaattccacca
 55 tggaaattttccatttgcaattatggctggtacaaccttagccctgttgttgacaggaac
 actgtattattaaagcctgctgaggatacagttttgactgcttataaattaatggaaata
 ttagaagaagcaggtttaccccaaggtgttgtaattttgttcctggtgatccaaaagaa
 attggagattatttagtcgaccataaagatacacattttgtcacatttacaggatcccgga
 gctacagggtacagctatttatgaacgtagtgctgtagtgcagaaggacaacagttttta

aaacgtgttattgcagagatgggtggcaagatgcgatagttgtagataataatgtagat
 acagatttagcggtgaagcaattgttacatctgcttttggtttctctggtcaaaaatgc
 tctgctgttctcgtgtcatagtcacatcaagacgtacatgatgaaatttgaaaaagca
 attcaattaactcaaaaattaacttttaggtaataactgaagagaacacatttatggggcca
 5 gtaattaatcaaaaacaatttgataaaaatcaaaaattatattgaaattggtaaaaaagaa
 ggcaaaactagagactgggtgggaacagatgattctaccggttatttcattgaaccaacg
 attttctccggactacaatctgcggatcgtatcatgcaagaagaaattttggaccagtc
 gtaggctttattaagggtcaaggattttgatgaggctattgaagtagctaatactgac
 tatggtttgacaggcgctgtaattactaatcatcgtgaacattggattaaggctgtgaat
 10 gaatttgatgtaggtaacctttacttgaatagaggtgtacagctgcagtagtggttat
 catccatttgggtgattcaagatgtctggtacagatgctaaaacaggaagtccagattac
 ttacttaatttcttagaacaataaagggttcttctgaaatgttttaa

Sequence 2682

15 MVVPFKNEPGIDFSVQTNVERFNEELRKVKAQLGQDIPLVINGEKLTKTDTFNSVNPANT
 Sqliakvskatqddiekafesanhayqswkwwshkdraelllrvaaiirrrkeeeisaimv
 yeagkpweavgdAAEGIDFIEYYARSMELADGKPVLDREGEHNRYFYKPIGTGVITIPV
 WNFPPFAINAGTTLAPVVAGNTVLLKPAEDTVLTAYKLMEILEEAGLPQGVVNFVPIDPKE
 IGDYLVLDHKDTHFVTFGTSRATGTRIYERSAVVQEGQFLKRVIAEMGGKDAIVVDNNVD
 20 TDLAEEAIVTSAFGFSGQKCSACSRVIVHQDVHDEILEKAIQLTQKLTGNTTEENTFMGP
 VINQKQFDKIKNYIEIGKKEGKLETGGGTDDSTGYFIEPTIFSGLSADRIMQEEIFGPV
 VGFIKVKDFDEAIEVANDTDYGLTGAVITNHREHWIKAVNEFDVGNLYLNRGCTAAVVG
 HPFGGFKMSGTDAKTGSPDYLLNFLEQKVSEMF*

25 Sequence 2683

Contig_0804_pos_874_1548,

is similar to (with p-value 3.0e-76)

>sp:sp|P39788|END3_BACSU PROBABLE ENDONUCLEASE III (EC 4.2.9
 9.18) (DNA-(APURINIC OR APYRIMIDINIC SITE) LYASE). >gp:gp|L4
 30 7709|BACYPIA_26 Bacillus subtilis (clone YAC15-6B) ypiABF ge
 nes, qcrABC genes, ypjABCDEFGHI genes, birA gene, panBCD gen
 es, dinG gene, ypmB gene, aspB gene, asnS gene, dnaD gene, n
 th gene and ypoC gene, complete cds's. NID: gll46223. >gp:gp
 |U11289|BSU11289_3 Bacillus subtilis 168 asparaginyl-tRNA sy
 35 nthesase (asnS) and endonuclease III (jooB) genes, partial c
 ds and DnaD protein (dnaD) and (jooC) genes, complete cds.
 NID: g533096. >gp:gp|Z99115|BSUB0012_174 Bacillus subtilis c
 omplete genome (section 12 of 21): from 2195541 to 2409320.
 NID: g2634478.

40 atggagagaatcctaataagtaagaaaaaagcattacaaatgattgacgttatagca
 gatatgtttcctaatagcagaatgcgaattaaaccatagaaatgattcgatcttacaata
 gctgtattattatcagcacagtgtactgataatctagtcaatcgtgtcactcaatcatta
 ttagaaaaatatcgaaacacgtgaagattattaaatgtgagtgatgaagaattacaaat
 gatatacgcctctattggattatatcgcaataaaagccaaaaatataaaaaaattatgccac
 45 tctttaattgaacaatttaattggtcaaatcccaaaacacataaagaattagagagtcta
 gctggagtgggcgtaaaacagcaaatgttgtaattgagtgctgcatttgagaaaccttct
 ttagctgtcgatactcatgttgagagagtttctaaacgtttgggaattaatcgttgga
 gatagtgtaagacaagtagaagatcgattatgtgatattatcccaagagatagatggaat
 aaaagccatcatcaattaatatttttgggagatatcattgtcttgctagaaaacctaaa
 50 tgtgagatatgtccgctgttaaatgattgtagagaaggacaaaaacgcataaagcaag
 ataaaggagcgctga

Sequence 2684

55 MERILMISKKKALQMIDVIADMFPNAECELNHRNAFDLTIIVLLSAQCTDNLVNRVTQSL
 FRKYRTPEDYLNVSDEELQNDIRSIGLYRNKAKNIKKLCHSLIEQFNGQIPQTHKELES
 AGVGRKTANVVMVAFGEPSLAVDTHVERVSKRLGINRWKDSVRQVEDRLCDIIPRDRWN
 KSHFQLIFFGRYHCLARKPKCEICPLLNDCREGQKRHKAKIKEA*

Sequence 2685

Contig_0804_pos_1553_1885,
 putative peptide of unknown function
 atgattgaaaaacaggatttcaatcatatagaggaccaacttgatcaactagcaagtaat
 aaacaactcaaaacaccagaagctagggaaacttttagatagttatttcgatttaattatt
 5 aattatttttaaaaaataaataacatagatgaaattcatttttaactcaactcgatacatat
 ccagtagttccaatgaattttgatgaacgctatcattatatggttgacgtaaacaccat
 tttatgggctatcgctcaaatgaaaacattgaaatcagaattaataaaaatgaatgcatt
 tatctaattagaaagcaacgtcaacaaaaataa

10 Sequence 2686
 MIEKQDFNHIEDQLDQLASNKQLKTPEARELLDSYFDLIINYFKQINNIDEIHFNQLDTY
 PVVPMNFDERYHYMVARHHFMYRQMKTLKSELIKMNASYLIRKQRQOK*

Sequence 2687

15 Contig_0804_pos_5054_0,
 is similar to (with p-value 1.0e-16)
 >gp:gp|AF076683|AF076683_2 Staphylococcus aureus oligopeptid
 e transporter putative substrate binding domain (opp-1A), ol
 igopeptide transporter putative membrane permease domain (op
 20 p-1B), oligopeptide transporter putative membrane permease d
 omain (opp-1C), oligopeptide transporter putative ATPase dom
 ain (opp-1D), and oligopeptide transporter putative ATPase d
 omain (opp-1F) genes, complete cds; and unknown gene. NID: g
 3800817.
 25 atgctcaaacgtacaattaaattcatactttatttaaatcgtaagttcgtttattatcttc
 atttttagttgagaagacatctggtaatccagcgattctgtatctacaacgtcatggttat
 acgtcgattacgcaagacaatattgaagcggcacaacatcaacttggttaggacaacat
 gtgttactaagatatatcgattgggttgacatgcactcacgggcaacttaggatacggc
 tttaagtacgaacgaagcagttaccgctatgataatggaagccatcggtccgacgc'tgtg
 30 ctaatcattgtctctagtgttatcatgttgccatttggttatattggttggttacttcgtt
 gggacgcgtccgcatacacgttacgctaattggaattcgtggattcgcccaagtgtatgacc
 tcaatgccagaataactggtagctatttttattcatttatttttaggcgtacgttggtgcaa
 ttggttaccatttgtaggtagtgattcatggcaacactttgtgctgccaatcttcacaatt
 gttgttatagaagggtgtcatatcttattgatgacagcacatctgattacacaaacgtta
 35 gatcaagatgcgtatcaactggcgagtttaagacatttttcgttaaaagcgcgtatcatc
 gtacaaattaaagagatatttgccacc

Sequence 2688

40 MLKRTIKFILYLIVSSFIIFILVEKTSGNPAILYLQRHGYTSITQDNIEAAQHQLGLGQH
 VLLRYIDWVGHALTGNLGYGFSTNEAVTAMIMEAIVPTLVLIIVSSCIMLPFGYIVGYFV
 GTRPHTRYANGIRGFAQVMTSMPEYWLAILFIYYLGVWRQLLPFVGSWSQHFVLPIFTI
 VVIEGCHILLMTAHLITQTLDDQDAYQLAQLRHFLSKARIIVQIKEIFAPX

Sequence 2689

45 Contig_0804_pos_3696_3097,
 is similar to (with p-value 5.0e-33)
 >gp:gp|AF068901|AF068901_4 Streptococcus pneumoniae penicill
 in-binding protein 2b (pbp2b), RecM (recM), D-Ala-D-Ala liga
 se (cdl), D-Ala-D-Ala adding enzyme (murF), MutT (mutT), cel
 50 l division protein FtsA (ftsA), cell division protein FtsZ (f
 tsZ), YlmE (ylmE), YlmF (ylmF), YlmG (ylmG), YlmH (ylmH), c
 ell division protein DivIVA (divIVA), and isoleucine-tRNA sy
 nthesase (ileS) genes, complete cds; and unknown gene. NID:
 g4009462.
 55 atgataaacgtaacattagagcaaatataaaactggatagattgtgaaattgatgaaaaa
 catttaaaaaaaacaataaatggcgtttcaattgattcacgaaaaatcaatgaagggcg
 ttatttataccttttaagggtgagaatgttgatggccatcggtttatcacacaagctttg
 aacgatggtgctggagctgttttttagtgaaaaagagaataaacattctgaagggaaacaa
 ggtcctattatttgggtagaagatactttaatagccttacagcaattggcaaaagcatat

5 ctaaatcctgttaaactcctaaggtgatagcggttactgggttctaattggaaaaacaaacaca
 aaagacatgattgaaagtgtattatcaactgaatttaaagttaagaaaacacaaggaaat
 tataataatgaaattggaatgccgttgactttactagaacttgatgaagacacagaaatt
 tctattctagaaatggggatgtcagggtttcatcaaataagagttgttatctcatatcgca
 caacctgatatagcgggtcatcacaacattggcgaatcacatatgcaagatttaaataa

Sequence 2690

10 MINVTLEQIKNWIDCEIDEKHLKKTINGVSIIDSRKINEGALFIPFKGENVDGHRFITQAL
 NDGAGAVFSEKENKHSEGNQGPIIWVEDTLIALQQLAKAYLNHVNPKVIAVTGSNGKTTT
 KDMIESVLSTEFVKVKTQGNYNNEIGMPLTLELDEDETEISILEMMSGFHHQIELLSHIA
 QPDIAVITNIGESHMQDLK*

Sequence 2691

15 Contig_0804_pos_2647_2078,
 is similar to (with p-value 3.0e-57)
 >gp:gp|Y17795|SAU17795_2 Staphylococcus aureus prfA, pbp2 ge
 nes. NID: g3955029.
 20 atggtaggttgcgcattagatatgaatactatcggttcggcttatggacatgggtgtgtca
 ggrgtaaacatgggtgcgaaaactgggtactggtagctatggacaagaaatatacga
 tataatttacctgataatgctgccaaagatgtttggattaatggtttcagtccagaatat
 actatgtcgggtatggatgggcttcaataagggttaacaaatggaacaaattcatttctc
 ggtcatccagaacaagattatccacaatacttgatgaagatgtgatgtctagtatctca
 tctaaagatgggtgaagatttcaaaaagcctaataatgatgtacaaggaagttcacccggacagt
 25 ctatctgtatcagggtcattctgataataataactactaaccgtagtggtcatggaagtagc
 gatacatcttcttcataaatgggtggctctaactcagcatcaagtggaaacaactcgaat
 agttcgaatggtaccagtcaaggtaactcaggcaatgcattttacacgtctgttcaattta
 aactctatatctcgattataaagtttcataa

Sequence 2692

30 MVGCRIRYEYRSAYGHGVSGVNMGAKTGTGTYGQEIYEKYNLPDAAKDVWINGFSPEY
 TMSVWMGFNKVKYGTNSFIGHSEQDYPQYLYEDVMSSISSKDGEDFKPNPDVQGSSPDS
 LSVSGHSDNNTNRSVHGSSDTSSSSNGGSNSASSGNNSSNGTSQGNSSNAFTRLNL
 NSIFDYKVS*

Sequence 2693

35 Contig_0804_pos_913_473,
 putative peptide of unknown function
 atgctttttcttacttatcattaggttctctccattcaaccagtcacaaatttttcagta
 40 gcttattttacaaaagatgttcacatgactgtttttcaatggatattctctgcatttact
 ttatggattggggttagtttatttttgacttttaggtattgattatcgctcaactcaacgat
 attcaaaaagcaagtagttttgccaatttacttaataattacactagctatattaggaggt
 ctatgggtttccagtatacacgtttcctgattggcttcagtcgattttctaaacacatgcc
 45 acatataatttaaagctacttgctatagatttagcgcaaaataaaggggtgaatatagaa
 gcgtttggctatctcggtgtctattgtataatctttgtgagtagttgtttattcatgaat
 aagaaaggagatgtacactaa

Sequence 2694

50 MLFSYLSLGFSPFNQSNFSAHFYKDVHMTVFQWIFSALTLWIGVSLFLTLGLIIAQLND
 IQKASSFANLLNITLAILGGLWFPVYTFPDWLQISIKHMPTYNLKLLAIDLAQNKGVNIE
 AFGYLVVYCIIFVSIALFMNKKGDVH*

Sequence 2695

55 Contig_0804_pos_0_452,
 putative peptide of unknown function
 gtgagaataggggaattaagtactttaatatatttaataatttccgatttttagctatattt
 gttgacaaacgtggtaattttcttaacttatttaattgtttgtactatttttataatcagc
 tatgtgacgatgatttatattttataataaccttagtgatagtagttttatattcattactg
 gttattcattatttaggaatcttttattttgtctatagtgcaatcctatgaatagattg

tttttcttctatagtgctttcgcccttaccttttatttttaaatgttcgtgtgtatcaaaa
gaatttataacctttttaatagctatgataagttgtttaataactatataatttaaat
ccaacatttgtggttccattaagtgcattttatttggttatattaattgttgctgtaggt
aattttaaaaatagagacgaacgaattatGAC

5

Sequence 2696

VRIGELSTLIYLIFPILAI FVDKRGNF LTYLIVCTIFIISVVTMIIFYKYLSDSLILYSL
VIHYLGIFYFVYSVNPMSLFFFYSAFALPFI FNVVRVSKEFITFLIAMISCLILTYIFN
PTFVVPLSAFYLVILIVAVGNFKNRDERIMT

10

Sequence 2697

Contig_0805_pos_575_1141,

putative peptide of unknown function

atgaaagggttaattattatagggagtgctcaagtaggggtctcatcacgaacgctttatca
aaatattttaaagggtcaactcggcgaacatgatgttgagggtgaaatctttgacctagct
gagaaacccattcatcaattggattttgctgggtacacacaagcagttgatgaaattaaa
aacaatgtcaaactctttacaaaataaagcaatggaagcagatttcttaattttaggaacg
ccaattatcatggatcggttttcagggtattcttaaaaatgcacttgaccaccttaatatg
gaccatttcaaaaatgaaacccgtgggactcatttgcataagtgagggaatagtaagttct
gagccattatcacacttgagagtcacgtacgtagtttacttggtattgctgtaccaacg
caaattgctacacatgattctgattatgctaaattagaagatggtaccttatacttagaa
gataatgaatttcaactacgttcaaaaattggttggatcaaattgtatccttcgtaaca
aatagtcacatatgaacacttaaaataa

25

Sequence 2698

MKGLIIIGSAQVGSHTNALS KYLKGQLGEHDVEVEIFDLAEKPIHQLDFA GTTQAVDEIK
NNVKS LQNKAMEADFLILGTPNYHGSFSGILKNALDHLNMDHFKMKPVGLICNSGGIVSS
EPLSHLRVIVRSLGLIAVPTQIATHDSYAKLEDGTLYLEDNEFQLRSKLFVDQIVSVFT
NSPYEHLK*

30

Sequence 2699

Contig_0805_pos_1585_2421,

is similar to (with p-value 1.0e-95)

>gp:gp|Y17116|SEY17116_1 Staphylococcus epidermidis gene enc
oding fibrinogen-binding protein, complete CDS. NID: g320154
9.

atgattaataaaaaaataattttactaactaaaaagaaacctatagcaaataaatccaat
aaatatgcaattagaaaattcacagtaggtacagcatctattgtaataggtgcaacatta
ttgtttgttttaggtcataatgaggccaaagctgaggagaatttcagtacaagactttaaa
gattcgaatacggatgatgaattatcagatagtaatgatcagtcagtgatgaagaagag
aatgatgtaattaataataatcagtcataaaactctgatgataataaccaataaaataaa
aaagaagaacgaataacaacgatggatatagaaaaagctcagaagatagaacagagtca
acaacaaatgtatagtaaaacgaagcaacatttttcaaaaagtcacctcaagataatact
catcttacagaagaagaggtaaaagaaccctcatcagtcgaatcctcaaattcatcaatt
gatactgcccacaacccatctcacacaacaataaatagagaagaatctgttcaaacaagt
gataatgtagaagattcacacgtatcagatttttgtaactctaaaataaaagagagtaac
actgaatctggtaaagaagagaatactatagagcaaccttaataaagtaaaagaagattca
acaacaagtcagccgtctggctatacaaatatagatgaaaaaatttcaaatacaagatgag
ttattaaatttaccataaatgaatataaagtaacgaaacttagcgtcactttcttcatt
gaaaagaaccgtgaaatacttttgactttcatatcaattctccttatgaattattaa

Sequence 2700

MINKKNNLLTKKKPIANKSNKYAIRKFTVGTASIVIGATLLFGLGHNEAKAEENSVQDFK
DSNTDDELSDSNDQSSDEEENDVINNNQSINSDDNNQINKKEETNNNDGIEKSSDRTES
TTNVDENEATFLQKSPQDNTHLTEEEVKEPSSVESSNSSIDTAQQPSHTTINREESVQTS
DNVEDSHVSDFANSKIKESNTESGKEENTIEQPNKVKEDSTTSQPSGYTNIDEKISNQDE
LLNLPINEYKVTKLSVTFIEKNREILLTFISILLMNY*

Sequence 2701

Contig_0805_pos_4398_4919,

putative peptide of unknown function

atgctcattgatatagttgttcttcttattatttggttactttatagtgatagggttccgt
 agaggtaatttggttatcgatattgcactttgcttcttcaattgtatctttatataatgcg
 5 tcacaacattatcaatctattgcgcaacgttttagttgtctttgtgccatttccgaaaacg
 gtggcgtttgacatggctctatactattccttatgatcatttgcaatacagatttgaaaaa
 gtgatagcattttattataatatttggatgtgtaagcttattttgtatctagttgttgtt
 acatttgataatataataacgtataaaaaagatacatttagtaagtcggatatcgagtgct
 gttttgagtatcatagcggtttttatatatattacaaattggactttatttattatcgcta
 10 tatccgcattcattttatacagtaaccaattatctcaatcgctattaagtcgagttgtgatt
 gaacaaattccttattttatcacaattttattttaatttataa

Sequence 2702

MLIDIVVLLIICYFIVIGFRRGIWLSILHFASSIVSLYIASQHYQSIAQRLVVFPFKT
 15 VAFDMVYTIPIYDHLQYRFEKVIAFIIIFGMCKLILYLVVVTFDNIITYKKIHLVSRISV
 VLSIIAVFIYLLQIGLYLLSLYPHSFIQYQLSQSLLSRVIEQIPYLSQFILNL*

Sequence 2703

Contig_0805_pos_4994_5461,

20 putative peptide of unknown function

atgacaaaaaaagatgtaattcaattattagaaaaaatagctatatatatatggagctaaaa
 ggagaaaaatacatttaaaagtttcagcgtatagaaaaagccgcacaaagcttagagggtgat
 gagcgtacatttagaagagattgatgatgtaacagaacttaaaaggcattggaaaaggcgta
 ggagaagttattaatgaattttaaacaacaaaggatcaatcatcgacccttcaagcacatcaa
 25 gatgaagtacctgaagggttagtgccacttttgaaaatacaaggattaggcataccattg
 ataattattatatttttcaacaattggatctatccacatatccatcgctttacaatctgg
 acaccagtcggcttcaaaacttgataattactggatcagtactaccaatagtagaattaaa
 ttcttcgggttgatgaatatattttcatcactatctactccttttgtaa

30 Sequence 2704

MTKKDVIQLLEKIAIYMEKLGENTFKVSAYRKAAQSLEVDERTLEEIDDVTELKGIGKV
 GEVINEFKTQGSSTLQALQDEVPEGLVPLLKIQGLIPLIIIFNNWIYPHIHRTIW
 TPVGFKLDNYWISTNSRIKFFGLMNIFHHYLLLL*

35 Sequence 2705

Contig_0805_pos_3875_2949,

is similar to (with p-value 2.0e-34)

>sp:sp|O07874|RNH2_STRPN RIBONUCLEASE, HII (EC 3.1.26.4) (RNA
 SE HII). >gp:gp|U93576|SPU93576_1 Streptococcus pneumoniae r
 40 ibonuclease HII (rnhB) gene, complete cds. NID: g2209338.

atgggaaatgtcgatatacaaaactcacgtcaaaagaaattcaatcattgatggctcaaact
 acttttgagacgacgaagttacctcaaggatgaaagctcgtagagatatcaaaat:act
 gttatcaatctctatagttctggcaaaagtaattgtttcaaggtaagaatgctgaac:actt
 gcgagtcgaattgctaccaaataaacaatcaacaactggcaaacatacatcatcaaatata
 45 actagtattcaatataatcgttttcattgtattggaagcgatgaagcaggcagtgggcgac
 tattttgggtccattgactgtatgtgcagcttatgtgagccaatcacatatcaaaatctta
 aaagaacttggtgtagatgattcaaaaaaactaagcgatactaaaatcgtagatcttgca
 gaacagctcattacctttatccgcattctttattaacattagataatgttaagtataac
 gaacgacaaagcttaggatggctctcaagttaaaatgaaagctgtcttacataatgaagct
 50 atcaaaaatgtgcttcaaaaaattgagcaagatcaactggattatattgttattgatcaa
 ttgtcaaaagcgagaagtttatcaacattatgcattatcagcattaccttttcttgacaaa
 acaaaatttgaaacaaaaggatgaatctaaatcactagcaatcgcggcagcaagcattatt
 tctcgttatgcatttgttaaacacatggaccacatctcaaaaaactccatattggaaata
 ccaaaaggagcaagtaacaaagtagatttaattgccgctaaagtcattcaaaaatgat
 55 attcaacaacttgatactatttcaaaaaaacatttttaaaaacagagataaagcaattcat
 cttatgaatcaaaaatacaataaataa

Sequence 2706

MGNVVYKLTSKEIQSLMAQTTFETTKLPQGMKARTRYQNTVINIYSSGKVMFQGNAEQL

ASQLLPNKQSTTGKHTSSNTTSIQYNRFHCIGSDEAGSGDYFGPLTVCAAYVSQSHIKIL
 KELGVDDSKKLSDTKIVDLAEQLITFI PHSLTLTDNVKYNERQSLGWSQVKMKAVLHNEA
 IKNVLQKIEQDQLDYIVIDQFAKREVYQHYALSALPFPDKTKFETKGESKSLAIAAASII
 SRYAFVKHMDHISKKLHMEIPKGASNKVDLIAAKVIQKYDIQQLDTISKKHFKNRDKAIH
 5 LMNQKYNK*

Sequence 2707

Contig_0806_pos_2408_3289,

is similar to (with p-value 0.0e+00)

10 >sp:sp|P37565|YACC_BACSU HYPOTHETICAL 31.8 KD PROTEIN IN FTS
 H-CYSK INTERGENIC REGION. >gp:gp|D26185|BAC180K_134 B. subti
 lis DNA, 180 kilobase region of replication origin. NID: g46
 7326. >gp:gp|Z99104|BSUB0001_71 Bacillus subtilis complete g
 enome (section 1 of 21): from 1 to 213080. NID: g2632267.
 15 atgaacatgattatagtgagaggttagcatcacggtgggaaataagagcatatgct
 gcaatcacacagagtcagtacagaagcacaacacgtcattatacatggcctactgct
 tctgccgctatgggaagaactatgacagctactgttatgatgggtgcaatgttaaaagga
 aaccaaagtttaacagttactgttgatggcaaaggtccaattggcagaattattgctgac
 gcagatgctcaaggaaatgttcgtgcatatgtagaccatccacaaacgcattttccactc
 20 aacgatcaaggtaaattggatgtacggcgagcagttggtactgatggttccattcaggtt
 gttaaagatgttggaatgaaagactacttttctggtgagagtcgaatagatcaggtgag
 ctaggagatgatttcacatactactatgccacaagtgaacaaacaccatcatcagtagga
 ttgggtgtattagtttaattccagacaactcaatcaaagcagcgaggagatttattattcaa
 gttatgccaggtgctactgatgaaacggtgactaaattagaagaagccattagtcgaatg
 25 caacctgtatcgaaattaattgagcaaggacttacacctgaaggaatattaaatgaaatt
 ttgggtgaaggtaattgttcaaattttaaatcaacgtcagcgcaatttgaaatgtaattgt
 agtcatgagaaatttttaaatgctattaaaggtttaggagaggcagaaattcatagcatg
 attaaagaggatcatggagctgaagctgtatgtcacttctgtggttaataaatatcagtat
 agtgaaagtgaattagaagattttattagaacaatgaaatag
 30

Sequence 2708

MTHDYIVRGLAYGGEIRAYAAITTESVQEAQTRHYTWPTASAAMGRMTATVMMGAMKLG
 NQKLTVTVDGKGPIGRRIADADAQGNVRAVVDHPQTHFPLNDQGLDVRRAVGTGDSIQV
 VKDVGMDYFSGASPIVSGELGDDFTYYATSEQTPSSVGLGVLVNPDSIKAAGGFIIQ
 35 VMFGATDETVTKLEEAISQMOPVSKLIEQGLTPEGILNEILGEGNVQILNSTSAQTECNC
 SHEKFLNAIKGLGEAEIHSMIKEDHGAEAVCHF CGNKYQYSESELEDLLETMK*

Sequence 2709

Contig_0806_pos_3617_4558,

is similar to (with p-value 0.0e+00)

40 >sp:sp|P37887|CYSK_BACSU CYSTEINE SYNTHASE (EC 4.2.99.8) (O-
 ACETYL SERINE SULFHYDRYLASE) (O-ACETYL SERINE (THIOL)-LYASE) (C
 SASE).
 gtgttttggatggcacaacacctgtagattatgttacacaaattattgggaatacacct
 gtagtcaaattaagaaacgttggtgatgatgcagctgatatttatgttaagttagaa
 45 tatcaaaatccaggtggttcggtaaaagatcgatcgcttttagcgatgattgaaaaagct
 gagegtgaagggaataaaacctggtgatacaatcggtgagcctacgagtggttaacact
 ggtataggtctagcatttgtatgtgctgccaaggggtacaaagcagtttttacaatgcct
 gaaacaatgagccaagagcgccgtaacttattaaaagcttatggtgctgaactagtatta
 50 acaccagatctgaagctatgaaaggtgcaataaaaaaagctaaagaattaaaagaagag
 cagggctattttgaaccacaacaattcgaaaatccagcaaatcctgaaattcatgaactt
 acaactggaccagaattagttgaacaatttgaaggtcgacaaattgatgcatttttagct
 ggtgtaggaactggtggtacgttatctggtgttggtaaagtattgaagaaagaatatcca
 aatgtggaatagtagctattgaacctgaagcttctccagttattaagcgggtggtgaacca
 55 ggccctcataaattacaaggattgggagcaggtttcgtacctgatactttaaatacagaa
 gtttatgacagcatcatcaaagtaggtaatgatactgctatggatattggcagctcgtgtt
 gctagagaagaaggtatattagcaggtatttcatctggtgctgcaatatatgctgctatt
 caaaaagcaaaagaattaggtaaaggtaaaacagttgtaacagattaccaagtaattggg
 gaacgttacttatcaacaccatttatattcatttgataattaa

Sequence 2710

VFWMAQKPVDYVTQIIIGNTPVVKLRNVVDDDAADIYVKLEYQNPGGSVKDRIALAMIEKA
 EREGKIKPGDTIVEPTSGNTGIGLAFVCAAKGYKAVFTMPETMSQERRNLLKAYGAELVL
 5 TPGSEAMKGAIKKAKELKEEHGYFEPQQFENPANPEIHELTTGPPELVEQFEGRQIDAFLA
 GVGTGGTSLSGVGKVLKKEYPNVEIVAIEPEASPVLSGGEPGPHKLQGLGAGFVPTLNT
 VYDSYIKVGNLTAMDMMARRVAREEGILAGISSGAAYAAIQKAKELGKGKTVVTVLPSNG
 ERYLSTPLYSDN*

10 Sequence 2711

Contig_0806_pos_5681_6160,

putative peptide of unknown function

gtgaaaacttataacaataatcaatctgtgggtaacttcttaagtgtttcagtacaagat
 ggtcaaacagtttaacaagggtgaacgtatcatcaattatgatacaaatgggaataaacgc
 15 caacaactattgaacaaagtgaatcaagcacaatctcaagttaatgatgattatcaaaaa
 gtaaatcaaagtcctaacaatcatcaattacaagttaaattgactcaagatcaaagtgt
 ttaaatgaagctcagcagtcattgtcacaaatgatgacagacaactcaatgacagcatgaat
 gcatcatttgatggtaaaattaacattaaaaatgattcagatgtaggcgaagggaacct
 attttgcaattaatttcttcaaatcctcaaatcaacgcaactatcacagagtttgatatt
 20 aataaaattaaagaaggcgatgaagtaaatgtcactgtaaatagcaaatgtcccatataa

Sequence 2712

VKTYNNNQSVGNFLSVSVQDQGVKQGERIINYDTNGNKRQQLLNKVNQAQSQVNDDYQK
 25 VNQSPNNHQLQVKLTQDQSALNEAQQSLSQYDRQLNDSMNASFDGKINIKNDSVGEQGP
 ILQLSSNPQINATITEFDINKIKEGDEVNVTVNSKCP1*

Sequence 2713

Contig_0806_pos_6623_6060,

putative peptide of unknown function

atgagaaaaggaaatcagaatgaagcttttagaagaatttatcggaactttattaaaagat
 gagcaatattattatgagtttagcatttttagaagtgaaacacaaaatcttgaaatcata
 atggagaagatgattaagcaaggaattacaaaatttcgtattgtacctttactcattttt
 agtgcattgcatttatcagtgatattccacaaataacttaagagatgaaagctcgatat
 35 ccacaaattgatagtaaaatgagtgccgctcttggtacacatccatatatgaaaacatta
 gtagaaaatagaattgctgatgaaaaagtcagtgaaagttcaaccaaagcaactatagta
 attgcccattggaatggaagtgacgttttacgaaagcacatgatgaattaaaagcattt
 gttaaaacgcttgatagtcacatcctgtttatgcaagagcttttatatgggacatttgct
 atttacagtgacatttacttcatcgcttctttaaattttattaatatcaaactctgtgat
 40 agttgcgttaatttgaggatttga

Sequence 2714

MRKGNQNEALEEFITLLKDEQYYYELAFLESETQNLIIIMEKMIKQGITKFRIVPLLI
 SAMHYISDIPQILKEMKARYPQIDSKMSAPLGTHPYMKTLVENRIADEKVSEGSTKATIV
 45 IAHNGSGRGFTKAHDELKAFVKTLDSHHPVYARALYGTFAIYSDIYFIAFFNFINKLCD
 SCVNLRI*

Sequence 2715

Contig_0806_pos_5264_4722,

putative peptide of unknown function

atgagtacacaggaattaaaagaaaatgatattcaaaactcaagtattccatttgtaaat
 cattttaacaaactaagacatcagccaaaatggatacttaaatcaatcatagtgatagta
 ttagctattattagcgctttcatcacatataatactagtaatgaactattagataaccaa
 tcgatagctaataagccaaatggatgaaaatattgtttcgatgtctacaactataggcgct
 55 ttcataaggacaactctttagtggttgtagtatttttaattcttctaattatttctaaa
 atattttaaatctgacgccaagcgagcagtttattctcagcagcactctcttattcaatc
 attatttttaggattttacaactattatttctttaaattcaaatagtttttggactaaagata
 acagactataagcttgatagcttaaacattttttcaaaggataataaaacactcatggac
 atctgctacttcatgatattcattttcaataagttttgagtatatattatcggcatat

taa

Sequence 2716

MSTQELKENDIQNSSIPFVNHFNKLHQPKWILKSIIVIVLAIISAFITYNTSNELLDNQ
 5 SIANSQMDENMFRMSTTIGAFIGTIFSVVVVFLIFLIISKIFKSDAKASSLFSAALSYSI
 IILGFTTIISLIQIVFGLKITDYKLDLNI FSKDNKTLMDICYFMISFISISFEYILSAY
 *

Sequence 2717

10 Contig_0806_pos_833_267,
 is similar to (with p-value 3.0e-54)
 >sp:sp|P37476|FTSH_BACSU CELL DIVISION PROTEIN FTSH HOMOLOG
 (EC 3.4.24.-). >gp:gp|D26185|BAC180K_132 B. subtilis DNA, 18
 0 kilobase region of replication origin. NID: g467326. >gp:g
 15 p|Z99104|BSUB0001_69 Bacillus subtilis complete genome (sect
 ion 1 of 21): from 1 to 213080. NID: g2632267.
 atgatacatggggcatttttcttagcattttcaaataaatcacgaacacgactcgcacca
 acaccancaaacaatctcaacgaagtcagatccactaattgagaagaatgggtgcaccagct
 tcacctgcaaccgcagctgctaataatgttttacctgtacctggaggcccaacaagtaag
 20 acaccttttggaattcttgaacccatttgtttaaatttcttgttatcttttaagaaatct
 acaatttctattaattcttgtttctcttcgtcagctcctgctacatctgagaaacgaact
 cgacgtttattactgtcgtacatcttagctttggattttccaaagttcatcatcacgacca
 ccaccgccgccaccttgggcttggctaaggaagaaaataataataatgcaatgattaat
 acaggaatcagtgctgtaaaaaactaacgaatacactttgttttcttcttctttaaact
 25 gtaaatttaagattatcttgtttcttagctgtatctgtgatttttgtaaatctttttca
 ttgttatataaaaattgtcgatgagtaa

Sequence 2718

MIHGAFFLAFSNKSRTRLAPTPTNISTKSDPLIEKNGAPASPATARANNVLPVPGGPTSK
 30 TPFGILEPICLNFLLSFKKSTISINSCFSSSAPATSEKRTRRLLSYILALDFPKFIIRP
 PPPPPWAWLRKKINNNAMINTGISVVKILTNTLCFSSSLTVNLRSLCFLAVSVIFCKSFS
 LLYKIVDE*

Sequence 2719

35 Contig_0808_pos_5584_4772,
 is similar to (with p-value 1.0e-62)
 >gr:gp|AJ223960|LLCAJ3960_4 Lactococcus lactis cremoris MG13
 63-inv1 chromosomal inversion junction DNA. NID: g3647234.
 atgaaagatcaattgagaggtaatgtcgtgattatacatcaacctgcagaagaaatgcc
 40 ccggggcgggtgcgaaggtatgatagatgacggcgtattagaaggggtggaccatgtttta
 ggtgcacatgtgatgagtatgatggaacaggtaaaatatactatcgtgaaggttttgtt
 caaacgggacgcgcttatttttagacttgttgtgaaaggtcaggggtggtcacggctcatcc
 ccacatacatcgaatgacgctattgttagcaggtgcgcattttgtaacgaccgcacagacc
 attgtttcaagacgcttaaatccgtttgaaacgggtgtagtcacaataggttccttcgac
 45 gggaaaggacaattcaatgtgattaaagatacaataacaatcgaaggtgatgtacgtgca
 ctgactgatgatacaagagacaatattcagaatgaaatgacacgactagtcagaggatta
 gaagagatgttcggagtgattttgtgattttgaatttaaaaaggattatccagctctttac
 aatgatcctgaatttacttcttatgtagcaacaactttaaaaaatgctaagctggatgac
 ataaaagcaatagatattttgtgagccgcagccaccctcagaagatttcgcgttttatgcg
 50 ttagaaaagacctcaacatttattttatcggtgcagcaccagaagatggacctatgtac
 cctcaccatcatcctaagtttaatatcaatgaaacatctatgctttagttgcagaggca
 gtgggtacaattgtattagattatttgaaataa

Sequence 2720

55 MKDQLRGNVVIHQP AEEMPPGAKGMIDDGVLEGVDHVLGAHVMSMETGKIYYREGFV
 QTGRAYFRLVVKGQGGHSSPHTSNDIAVAGAHFVTTAQTIVSRRLNPFETGVVTIGSFD
 GKQ-QFNVIKDTITIEGDVRLTDDTRDNIQNEMTRLVRGLEEMFGVICDFEFKKDX3ALY
 NDPEITSIVATTLKNAKLDDIK AIDICEPQPPSEDFAFYALERPSTFIYSGAAPEDGPMY
 PHHHPKENINETSMLVVAEAVGTIVLDYLK*

Sequence 2721

Contig_0808_pos_4094_3633,

is similar to (with p-value 3.0e-18)

- 5 >sp:sp|P38049|YIXC_BACSU HYPOTHETICAL 18.8 KD PROTEIN IN PBP
F 5'REGION. >pir:pir|B40614|B40614 hypothetical protein X (p
bpF 5' region) - Bacillus subtilis >gp:gp|L10630|BACBPBF_2 B
acillus subtilis penicillin-binding protein (pbpF) gene, 5'
10 end. NID: g304158. >gp:gp|Z99109|BSUB0006_86 Bacillus subtil
is complete genome (section 6 of 21): from 999501 to 1259940
. NID: g2633260. >gp:gp|Y14083|BSY14083_4 Bacillus subtilis
chromosomal DNA, region 76-78 degrees: between glyB-aprE. NI
D: g2226224.
atgaagggtatgtcttatgttaattagcgatacaataaaacatatataatagaagaaact
15 agtaattcattcacaaattgaaaagaataatgatcagcagcactacgaagtattagaatca
atcaacagcttatctaattgattcattttgtgtgttaaatcattattcgtcaatggaggt
aatgaagaggtttttgagtcacggttttttaagcgaaatcaacatttgcaagatgtgcct
ggttttaagcggttaagatttcttagaccggtagtcaaagggagacattacattatcatc
acgctatggaacagtagacaagctttctatgattggcaaaattcacaagcatatgcgcaa
20 actcataaaaaaacgtggaactcaaaaaggtgttgatcatcgtatagtcattagagattta
tcctataatataagaatagaggttagaaagtcttaataactaa

Sequence 2722

MKGCLMLISDTNKTYIIIEETSNSFTIEKNNDQQHYEVLESINSLSNDSFCVLNHLFVNGG
25 NEEVFESRFLKRNQHLQDVPGFKALRFLRPVVKGRHYIIITLWNSRQAFYDWQNSQAYAQ
THKKRGTKQGVDRIVNRDLSYNIRIELESLNN*

Sequence 2723

Contig_0808_pos_1986_1399,

is similar to (with p-value 4.0e-51)

- 30 >sp:sp|P37470|SP5C_BACSU PROBABLE PEPTIDYL-TRNA HYDROLASE (E
C 3.1.1.29) (PTH) (STAGE V SPORULATION PROTEIN C). >gp:gp|D2
6185|BAC180K_116 B. subtilis DNA, 180 kilobase region of rep
lication origin. NID: g467326. >gp:gp|Z99104|BSUB0001_53 Bac
35 illus subtilis complete genome (section 1 of 21): from 1 to
213080. NID: g2632267.
gtggaggtaacaataatgaaatgcattgtcggctcttggaacattggtaaactgtttgaa
ttaacaagacataatattggtttcgaagttgtcgatgatattctagaacgccaccaatatt
actttagacaaaacaaaaatttaaaaggtgcataactattgaacgttttaacggcgaaaaa
40 gtattattttattgagccaatgacctgatgaacttatctggtcaagctgtagcccttta
atggattattataatgtcgatgttgaagatttgatcggttttatatgacgatttagattta
gaacaaggacaagtgcgtctgcgcaaaaagggagtgaggcggtcataatggtatgaaa
tcgataattaaaatgcttggtacagatcaatttaaacgtattcgaattggtgttgccgt
ccaacaaatgggatgtctgttccgactatgttttcaaaaaattttcaaaaagaagaaatg
45 atcattatggaaaaggttaattgaacattctgcaagagctgtagaatcttttattgaaagt
tctcgttttgatcatgttatgaatgaatttaattggtgaagtcaagtga

Sequence 2724

VEVTIMKCIIVGLGNIGKRFELTRHNIGFEVDDILRHQFTLDKQKFKGAYTIERVNGEK
50 VLFIEPMTMMNLSGQAVAPLMDYYNVVDVDELIVLYDDLLEQQQVRLRQKGSAGGHNGMK
SIKMLGTDQFKRIRIGVGRPTNGMSVPDYVLQKFSKEEMIIMEKVIEHSARAVESFIES
SRFDHVMNEFNGEVK*

Sequence 2725

Contig_0808_pos_0_1393,

is similar to (with p-value 4.0e-65)

- 55 >gp:gp|AF054624|AF054624_1 Lactobacillus sakei transcription
-repair coupling factor (mfd) gene, partial cds; L-lactate d
ehydrogenase (ldhL) gene, complete cds; and unknown genes. N

ID. 35.1014.

atgattgcaaattacattagcgaagataatcgttttcaagaattagatgaagtccttcggc
 caagagaatatttttagttacgggattatccccgtcagcgaaggcaacaattatgctgaa
 aaatattttaaagatcataaacaatgctactcgtaactaataatttataccaagcagat
 5 aaaatcgaaactgacattttacaatatgtagatgactcagaagtttataaatatcctggt
 caagataataatgactgaagagtttttctactcaaagtcacaaattgatgagtgagcgtggt
 agaacggtgactgcttagcccaaggcgaaaaagggttatttattgtgcctttaaacggc
 tttaaaaaatggctaacacgggttgatttatggaaagatcatcaaagacgcttaaagta
 ggtcaggatattgatgttgatgcattcttaataaattagttaatatgggttatcgccga
 10 gaaagtgtagtgctacatatattgggtgagttttcattgctggtggaatcatagatatatat
 ccggttgataggaacacctgtgagaatagagctatttgatactgaagttgattccatcaga
 gactttgatgtagaacaacacggttctaacgataatataaatcaagttgaaatcacaca
 gctagtactacattattactgatgaagtgatacaacacttacaaaatgaacttaaaaaa
 gcatatgaatatacacgcccataaaattgaaaagtcgcgtacgtaattgatttaaaagagaca
 15 tatgaaagttttaagttgtttgaatctaccttttcgatcatcaattgttacgacgtctt
 gtttcatttatgtatgaaaaacacaccccttattgactattttcaaaaaacgcgatt
 attgtagttgtagtttaactcgatttaaggaaacagaagaaacacttacacagaagtt
 gaagattttatgagtaacttaattgagagtggaatggatttatcggaacaggatttatg
 aagtatgaaagttttgacgcattattagagcaacatgccgttgcatatttcacattattt
 20 acctcttcgatgcaagtaccattacaacatatatttaagttctcttgtaaacaggttcaa
 caattttatggtcaatatgacattatgcgctcggaatttcaaagatacgtgcatcagat
 taaactgtcgtagttcttgttgaaactgaaacaaaagttgaacgtattcaatcaatgctt
 aatgaaatgcatattccaacagtatcaaatattcacgaagatattgatggtggtcaagtt
 gtatgacggaagtgagttcttgaaggccttgaattaccttatatgcagttggtagtc
 25 atcacagaaagag

Sequence 2726

MIANYISEDNRFOELDEVFGQENILVTGLSPSAKATIIAEKYLKDHQMLLVNNLYQAD
 KIETDILQYVDDSEVYKYPVQDIMTEEFSTQSPQLMSERVRLTLALAQGEKGLFIVPLNG
 30 FKKWLTPEVDLWKDHQMTLKVQDIDVDAFLNKLNVNMGYRRESVSVSHIGEFSLRGGIIDIY
 PLIGTPVRIELFDTEVDSIRDFDVETQRSNDNINQVEITTASDYIITDEVIQHLQNELKK
 AYEYTRPKIEKSVRNDLKETYESFKLFESTFFDHQLRLRLVSMFYKPKSTLIDYFQKNAI
 IVVDFENRIKETETLTTEVEDFMSNLIESGNFGFIGQGMKYESFDALLEQHAVAYFTLF
 TSSMQVPLQHIKFSCKPVQQFYGYQYDIMERSEFQRYVHQDYTVVVLVETETKVERIQSML
 35 NEMHIPTVSNIHEDIDGGQVVVTEGSLSEGFELPYMQLVVITERX

Sequence 2727

Contig_0810_pos_182_592,

putative peptide of unknown function

40 atggaagacgaatgcgcaaaactgaaagcttggttatcaataatcctgtgttactcaat
 gaacatgaagacgaagcagatatatctgtatataggatttatatctactaaaggtgctatt
 ggagaaggtgcagaaagactagaacgacatggtgtaaaagtgaatacgtatgcattatcga
 caattacatcctttccctaaagatatgttcaacaagctattaataaagcttcgaaagta
 atagttgcagaaacataattatcaaggacaattatcaagtattttaaaaatgaacacacaa
 45 gttaatgataaattagttaatcaaacaaaatacgtatgggaaacctttcttaccttatgaa
 attgaagaaaaaggtttggaaattgctaaagagttaaaggagttggtgtaa

Sequence 2728

MEKMRKTESLVINNPVLLNEHEDEADILYIGFISTKGAIGEGAERLERHGVKVNMTMHIR
 50 QLHFFPKDIVQKAINKASKVIVAENHYQGQLSSILKMNTQVNDKLVNQTKYDGKPFPLPYE
 IEEKLEIAKELKELV*

Sequence 2729

Contig_0810_pos_2045_3106,

putative peptide of unknown function

55 atggacaaattttaaattctatgacagaattaaaagaattgactaaagaaggaaaagattgg
 gaaatagagtgtaaaatcggttctagcatagtcactatattagcattacatggcggtgga
 attgaacctgccacaactgaattagcctatacaattgcacattgtggcgactataactat
 ttttcctttaaaggtatgagaagtaaggggaataatgagttacatgtgacttccacacat

tatgatgaccaaattgcattagatttagtgagaggtagccaaagaactgtagccatccat
 ggttgtagaggaatgaaagtgtggcttatataggaggtagtgatgacagactaattgag
 ttaatcaccgaatctcttgaagatataggaattagcgtgcgagaagcaccacatcatatt
 tctggaactcaagaaaatactgaaaatggagcagtcattctataaaggtaataactttac
 5 aattatcgaagttttgatcagtatatttttcaaaaagttgtaaattatttaaatttgaat
 caaaagataaacaatttgattatttgggtgtaaaaagtgcatatattttaaaagaaaca
 agcgaagcatttaagcaagatgcacgtacatatattcaccaactaatagaggttgactcc
 ttacaaacattacctgatgatgattatgtgaaaattgctttcaatataaatcgtcagact
 catccagacttagatgagaaattagctcttaagtttaagacgatattaaactagtatca
 10 agtgggagagatagtatagatgttattatgccaaatatgactaagggtcaagctttgtct
 agattattaaaagaatggcaaatgcctgcttcacatttaattggcatttgagatgcaaat
 aacgataaagatatgttggagcttgccgaacatagttatgttatggctaatagtgaagat
 caatcattatttaatatagcgagtcattgtggcaccttccaatgatgaacaaggcgacta
 tcaacaatcgaaaatgttgttctcggttattccaataaataa

15 Sequence 2730
 MDKFKSMTELKELTKEGKDWEIECENRSSIVTILALHGGGIEPATTELAYTIAHCGDYN
 FSFKGMRSGKNNELHVTSTHYDDQIALDLVRGSQRTVAIHGCEGNESVAYIGGSDDR
 LIE LITESLEDIGISVREAPHHISGTQENTENGAVIYKGNQLYNYRSFDQYIFQKV
 VNYLNLN 20 QKINNLICGVKSAYILKETSEAFKQDARTYHQLIEVDSLQTLPDDDYVKIAFNIN
 RQT HPDLDEKLALKFKDDIKLVSSGRDSIDVIMPNMTKGQALSRLLEWQMPASHLMA
 FGDAN NDKDMLELAHSYVMANSEDQSLFNIASHVAPSNDEQGVLTSTIENVVLGY
 SNK*

Sequence 2731
 25 Contig_0810_pos_4526_5194,
 is similar to (with p-value 2.0e-20)
 >gp:gp|AF012552|AF012552_2 Helicobacter pylori prolipoprotei
 n diacylglycerol transferase (lgt) and NADPH-linked flavin n
 itroreductase (rdxA) genes, complete cds. NID: g2564440.
 30 atgattatgaatcagatgaatcaaacgattattgatgcattccattttagacatgcgaca
 aaagaatttgaccctacgaaaaaatttagtgatgaagattttaatacgaattttagaaca
 ggtagatctctccaagttcactaggtttagaaccttggcactttgtagtgttcwaaat
 aaagaattgagagaaaaattgaaagcctatagttggggagcaciaaagcaacttgataca
 gcaagtcactttgtattaatttttgcctgtaagaatgtgacggctcatacagattacgtg
 35 caacatttacttcgtggcgtcaaaaaatgaagaaagtacaattccagcagttgaaaat
 aaatttgatgatttccaagaaagtttccatattgccgataatgaacgaacattatatgac
 tggcgagtaaacaaacatatattgcattagcaaacatgatgacaagtgctgcattacta
 ggtatcgactcatgtccaattgaaggatttgatttagataaagtgactgaaattctttca
 gatgaggggtgttttagatacggacaatttgggtatttcagttatggtaggctttggttac
 40 agagcacaagaacctaaacatggcaaaagttagacaaaacgaagacgacatcattagttgg
 attgaataa

Sequence 2732
 45 MIMNQMNQTIIDAFHFRHATKEFDPTKKISDEDFNTILETGRLSPSSLGLEPWHFV
 VVQN KELREKLKAYSWGAQKQLDTASHFVLIFARKNVTATHTDYVQHLLRGVKKYEESTI
 PAVEN KFDDFQESFHIADNERTLYDWASKQTYIALANMMTSAALLGIDSCPIEGFDL
 DKVTEILS DEGLDTEQFGISVMVGFYRAQEPKHGKVRQNEDDIISWIE*

Sequence 2733
 50 Contig_0810_pos_4292_3300,
 is similar to (with p-value 0.0e+00)
 >gp:gp|U31175|SAU31175_1 Staphylococcus aureus D-specific D-
 2-hydroxyacid dehydrogenase (ddh) gene, complete cds. NID: g
 1644432.
 55 atgacaaaaattatgtttttcggcacaagagcatatgagaaggacatggcattacgttgg
 ggaaagaaaaataatcgcattgacacacacagaacttttaagtgtagatactgtc
 gatcaattaaaagattatgacggtgttacaacaatgcagttcggttaattagaacctgaa
 gtttaccctaaattagatgcctatgttattaaacaaattgcacaacgtacggctggatt
 gatattgatgacttagaacttgcaaaaaacatgaaattattatctcgaatatacctagt

tattcacctgaaacaattgctgaatattcggtatctatcgctctgcaactcgtagcaaaa
 ttcccaacaattgaaaaacgtgtgcaagcacataatttcacatgggcgtcccctattatg
 tctcgtccagtaaaaaatgatgactgtagcaatcatcggtacagggcggtattggtgcgca
 actggtaaaatctatgctggttttgggtgcgagagtagttggttatgatgcataatc:taat
 5 cattctttatctttcttagaatataaagaaacagtagaggatgcattaaagatgctgat
 attatctcattacatgtacccgctaataaagatagtttccatttatttgataacaatatg
 tttaaaaatggttaaaaaagggtccggttttagtcaatgccgcaagaggagctgtgataaac
 acgcctgatttaattgaagcagtaataatggtacattatcaggtgctgccattgacaca
 tatgaaaatgaagctaattatttcacatttgattgttcaaatcaaacgattgacgaccca
 10 atattattagacctaatgaaaatgaaaatattttagttacacctcatattgcctttttc
 tccgatgaagcagtagcaaaaatttagtagagggtggtttgaatgcagcattatcagtaatt
 aatactggcacatgtgatacgcgattaaactaa

Sequence 2734

15 MTKIMFFGTRAYEKDMLRWGKKNNIDVTTSTELLSVDTVQDKDYDGVTTMQFGKLEPE
 VYPKLESYGIKQIAQRTAGFDMYDLELAKKHEIISNIPSYSPETIAEYSVSIALQLVRK
 FPTIEKRQVQAHNFTWASPIMSRPVKNMTVAIIGTRIGAATGKIYAGFGARVVGDAYPN
 HSLSFLEYKETVEDAIKDADIISLHVPANKDSFHLFDNNMFKNVKKGAVLVNAARGAVIN
 TPDLEAVNNGTSLGAAIDTYENEANYFTFDCSNQITIDDPILLDLIRNENILVTPHIAFF
 20 SDEAVQNLVEGGLNAALSVINTGTCDTRLN*

Sequence 2735

Contig_0812_pos_5666_6487,
 is similar to (with p-value 1.0e-57)
 25 >gp:gp|U75480|SMU75480_1 Streptococcus mutans putative HPr(s
 er) kinase (ptsK) and putative prolipoprotein diacylglycerol
 transferase (lgt) genes, complete cds. NID: g3924622.
 atggctggatattttcacattatgcttcagaccgtattcaattattagggacaacggag
 ttatcattttataatttacttccagatgaagagaagaaaggaagaatgagaaaattatgc
 30 cgacctgaaactccagcgattattgttacacgtgggttagaaccacccgaagaacttata
 caagcatctcaagaaacgcatacaccaattattgttgcaagatgccacaacgagttta
 atgagtaggttaacgacatttctcgaacatgaactcgcaaaaactacttctttgcacggt
 gtacttgttgatgtttacggtgtaggtgactaattacaggagattctggcattgggaaa
 agtgaactgcattagaattagtcacacgaggccatagattagtggtgatgataatgta
 35 gaaatcaagaaattactaaggatgaacttgtagggaaaccgcctaaacttatcgaacat
 ttgctagagattcgtggtctcggaatcattaatgttatgactttgtttggagcaggatca
 atattaactgaaaaacaagttcgattaataatatttagaaaaattggaataagaataaaa
 ttatacgcgctgtaggtcttaatgaagaaacattgaaaattcttgatacggaaatcact
 aaaaaaacgataaccagtttagaccagggcgtaatgtagcagtaattattgaagtagctgct
 40 atgaattatcgctcttaatatcatgggtattaatacagcagttgaatttaatgagagactt
 aatgaagaaatcggttcgaaatagtcataaaagttagaggagtaa

Sequence 2736

45 MAGYFSKYASDRIQLLGTTTELSFYNLLPDEEKKGRMRKLCRPETPAIIVTRGLEPPEELI
 QASQETHTPPIIVAKDATTSLMSRLTTFLEHELAKTSLHGVLDVYGVGVLTGDSGIGK
 SETALELVKRGHRLVADDNVEIKEITKDELVGKPKLIEHLEIRGLGIINVMTLFGAGS
 ILTEKQVRLNINLENWNKNKLYDRVGLNEETLKILDTEITKKTIPVRPGRNVAVIEVAA
 MNYRLNIMGINTAVEFNERLNNEEIVRNSHKSEE*

Sequence 2737

50 Contig_0812_pos_6493_0,
 is similar to (with p-value 8.0e-40)
 >sp:sp|P52282|LGT_STAAU PROLIPOPROTEIN DIACYLGLYCERYL TRANSF
 ERASE (EC 2.4.99.-). >gp:gp|U35773|SAU35773_1 Staphylococcus
 55 aureus prolipoprotein diacylglycerol transferase (lgt) gene
 , complete cds. NID: g1016769.
 atgaatataacattagatatcgatcctgttgccttttagcttgggaccaatccaagtt
 cgatgggtatggaattattattgcttgtggtatcttacttggatcttcattgcacaagca
 gcattgaaacaggttggttacataaagacaccttaatcgatattatattttatagcgcg

attgttggaattcatagttgcgagaatatactttgttacatttcaatggccatattacatg
aatcacttgagtgagataccaaaaatttggcatgggtggtattgccatacatgggtggctta
attgggtgacttatctctgggattattgtttgtaaaatcaaaaatctacatccgtttcaa
ataggagatattgtggc

5

Sequence 2738

MNITLGYIDPVAFLGPIQVRWYGIIACGILLGYFIAQAALKQVGLHKDTLIDIIFYSA
IVGFIVARIYFVTFQWPYYMNLSEIPKIWHGGIAIHGGLIGGLISGIIVCKIKNLHPFQ
IGDIVA

10

Sequence 2739

Contig_0812_pos_6327_5983,

is similar to (with p-value 3.0e-31)

>gp:gp|U75480|SMU75480_1 Streptococcus mutans putative HPr(s
er, kinase (ptsK) and putative prolipoprotein diacylglycerol
transferase (lgt) genes, complete cds. NID: g3924622.

15

gtgatttccgtatcaagaattttcaatgtttcttcattaagacctacacgatcgataat
ttattcttattccaattttctaaattaatatttaaatcgaactgtttttcagttaatatt
gatcctgctccaaacaaagtcataacattaatgattccgagaccacgaatctctagcaaa
20 tgttcgataaagtttaggcggtttccctacaagttcatccttagtaatttctttgatttct
acattatcatcagccactaatctatggcctcggttgactaattctaatagcagtttcaact
ttcccaatgccagaatctcctgtaatttagtacacctacaccgtaa

20

Sequence 2740

25

VISVSRIENVSSLRPTRSYNLFQFSKLIENRTCFSVNIDPAPNKVITLMIPRPRISSK
CSISLGFGPTSSSLVISLISTLSSATNLWPRLTNSNAVSLFPMPESPVISTPTP*

Sequence 2741

Contig_0813_pos_1018_2097,

putative peptide of unknown function

30

atgcttgaaaaaacattcgaagtcacgtatacaaatgaacaaaaaattgaattagaagca
caattgttttcaacacaacttttatttcaatttctcttttcgcaaggtaggttagaagaa
gcccgaacatatattttgaatcaatcttacgagatacaacagcatagggtgattaggaat
ttacttgcaatgtgttatttgtatctaggtgagtagatgatagcgccaaagcaatgtttgaa
35 gaacttttaaaggaagataattcagacgtgcacgttctcactacacattattactt
tataataaaaaagaaacagaaaaatatcaaaaatatcttaaaatacttaataaagtagta
ccactaaatgacgacgaaacccctttaaattaggaatcgattgagttattttaaacagtat
cgtgcttctcaaaatttactttatccactttataaaaaaggtaaatttgtctctattcaa
atgtataatgcattgagtttcaatttttattacctaggaaataaagacgaaagtattgag
40 atgtggaacaagctcactcaaatcttgaagttgatgttggttatgcaccttgggtaatt
gaggaaagtaaaacggtattttgaatcacgagtggtaccattattactagatgataataat
cattatcgactttacggtatttttttacttcatcaattaaatggaaaagaaataactaatg
actgaagatatatttggtcaattcttgaatcaatgaatgactatgagaaactttatctcaca
tatttgggtacaaggactcacactcaataaattagattttatacacagaggtagtcaaaagg
45 ttgtataattttaagaaattcaaatataacacgctctttatttacagattggattaatcaa
gcagaaatgattatagctgaaaatgtagatttagtagatgtcgatagatatgtagctgca
tttgtttacctatcgatcgctgcttagccaaccactaccaagaggcaattgatggac
gattttaatgtttctagatacaaaactgaataaagcaattgaatttatattgagcatataa

50

Sequence 2742

MLEKTFEVTTYTNEQKIELEAQLFSTQLLFQFLFSQGRLEEARTYILNQSYEIQQHRVIRN
LLAMCYLYLGEYDSAKAMFEELLKEDNSDVHALCHYTLNKKETEKYQKYLKILNKVV
PLNDDETFKLGIVLSYKQYRASQNLPLYKKGKFVSIQMYNALSFNFYLGKNDSEIE
55 MWNKLTCTSEVDVGYAPWVIEESKTVFESRVLPDLLDDNNHYRLYGIFLLHQLNGFEILM
TEDIWSILESMDYDEKLYLTYLVQGLTLNKLDFIHRGMQRLYNFKKFKYNTSLFTDWINQ
AEMIIAENVLDVVDRIYVAAFVYLSYRRSSQPLTKRQLMDDFNVSRYKLNKAIEFILSI*

Sequence 2743

Contig_0813_pos_2162_2497,

is similar to (with p-value 6.0e-50)

>gp:gp|AJ223781|SAAJ3781_1 Staphylococcus aureus trxB gene.

5 NID: g3582102.

atgactgaagtagattttgatgtagcaataatcgggtgcaggtcctgccggtatgacagca
gcagtatatgcatctcgtgccaatTTAAAAactgtcatgattgaacgcggtatgc.aggc
ggTcaaatggcaaacactgaagaagtagagaattttccaggatttgagatgacacaggt
cctgacttatctactaaaatgtttgaacatgctaaaaaatttggtgcggaataccaatat
10 ggcgatattaaatctgttgaagataaaggcgactataaagttatcaatttagggaatata
gttgttgatgtaaatgtgtgtcacagtattgttta

Sequence 2744

MTEVDFDVAIIIGAGPAGMTAAVYASRANLKTVMIERGMPGGQMANTEEVENFPGFEMITG
15 PDLSTKMFHAKKFGAEYQYGDIKSVEDKGDYKVINLGNIVVDVNVCHSIV*

Sequence 2745

Contig_0813_pos_5893_6465,

putative peptide of unknown function

20 atgaaaaaagtatttagcactattatttgcgtcaacactcatttttaggagcatgtggggac
aaaaatgacgaatctaaaaatgattcttcatctaatctacagatagcggtttctgtagac
aaaaatgacaatgaagataaaaacaacagcgaattcaaaaatgataaattaacaacagac
aattttgatattgaaattttagaagccaaaactgtcaaagcatctgagtagcatgatgac
aagaaaccaagtatcgctatcatctatggtgttaaaaaataaaaaagacaaagatttaaca
25 gcatcttcagcatttatcgaaatcatttgatatttatcaaaattctaaagatgttaagaga
agattagaaattggcgggtggatgatgatactgatttaaaaggaaaaatatgaagaagatg
gatacaaaaattaacaaagatggaaaaagtaaagggtgttatgttctttaattgaaagac
actaaacaccagtaacacttgaagctaaagatccaaatcatagcaataacgaaaaagt
30 ggtactaaagaatttaaaattaaagaaaaataa

Sequence 2746

MKKVLALLFASTLILGACGDKNDESKNDSSNSTDSVSVVDKNDNEDKTTAKFKNDKLTTD
NFDIEILEAKTVKASEYDDDKKPSIAIIYGVKNKKDKDLTASSAFIESFDIYQNSKDVKR
RLEIGGGYDTDLKEKEYEEDMDNKINKDKGKVGVMFFKLKDTKTPVTLEAKDPNHSNNEKV
35 GTKEFKIKEK*

Sequence 2747

Contig_0813_pos_6510_0,

putative peptide of unknown function

40 atgtctcaaaaaattaaagttatagtgcctatcgacttggattaatcattcttttgggg
attgcttgggggtgatatgcctttgtaacaaatacacctaaaaatgcatatttactgagc
gaaaaaaagaccgctataaatgtaaaatcttatgttgatcatcggttttagtaacgaaaag
aaattccaaaaaaaattaaaagataattcatatgttaatacgtataatctacatgcta
gcatctaaggaatatctaaaagatcttggtttacctaaaactattttagatagttctaaa
45 ataactggaactatcggtcatgatccaaaatcaaataaaggaatcatgagcgtatcacct
aaaatattagataaagatattggtaagttccaatggacagcaaatgattcaactcaattc
ttcgaatcaccttattcaagaaaaagtatagcgtcaaaaactcagaactattagaaca
gctgctcaaatctttgatgaagatccttctgactataaagaagagggactttcaaatgca
aactttgatctgaataataaaattgggtattgttcattctcaacaagaagatgttaaaaa
50 ctgattaaagcgatatacagatttagtcacatgataaattagaagatgatgactttgaaaa
ggtaaaaaagaaaaagttaaaattgacgggtgaacaaaaaaatttaaacctatcacttta
aatataagtcgtgataaagcaaaaaaaatcactgtcgcagcattaaaaaaagctaaaaat
gataaagaattac

55 Sequence 2748

MSQKIKVIVPIVLGLIILLGIAWGVYAFVTNTPKNAYLLSEKKTAINVKSYVDHRFSNEK
KFQKKLKDNSYVNTYNLHANASKEYLKDGLPKTILDSKITGTIGHDPKSNKGIMSVSP
KILDKDIGKFQWTANDSTQFFESPLFKKKYSVKNSELLETAQIFDEDPDSDYKEEGLSNA
NFDLNNKLGIVHSQQEDVKKLIKRYTDLVIDQLEDDDFEKGKKEKVKIDGETKNLKPITL

NISRDKAKKITVAALKKAKNDKELX

Sequence 2749

Contig_0813_pos_5275_4484,

5 putative peptide of unknown function

atgacaattcttgttcattcaaagcatcaaccgagcgaatacgcagcgattgcacatcag
 ttgatggcgacaacacatgtgtgtgtgaacaagtagggttcattgaatcagtaaactat
 gaaaatggggataactatcacttggtaatgagtggaatgaattttgtggtaatgcgact
 atgtcttacattcactattttaaagaacgtttattgatacagcatcaacaatttcaatta
 10 agagtttcgggatgttctcatcctgtagagtgtaaagttcattcgcaacattatgaagt
 actatgccaaaagtacatcaagttaaggaaagatttgtgaaattaggggaccaacagttt
 aaagcatttgaaattagatacagatacatacattcattacgtgttgatgtgtgatggtgta
 gatttagcaatgaaacagcgcggtggaagatttgtcagtcgcgcaacatggcatcaacaa
 tttaaaacgattggcgatcatgctttttcaacaagataaacaattcatatatccactgata
 15 catatacctaaaaatagatagcttaactctgggaaaatagctgtggttcaggagcggttct
 atcgggtgtgttagttaattatctaacagatcatgatattcaagattacctaagtaaccaa
 ccgggagcgagtattattgtctcatccagaaagtctggacaaaatgaatacacaacacg
 attaagtgtcaagtttcaactgtcgcaacaggacaagcatatatagaacaggagacaatg
 acgcaaatatga

20

Sequence 2750

MTILVHSHKQPSEYAAIAHQLMATTHVCCEQVGFIESVNYENGDNHYHLVMSGNEFCGNAT
 MSYIHYLKERLLIQHQFQLRVSGCSHPVECKVHSQHYEVTMPKVHQQKERVKLGDQQF
 KAFEIRYDITYIHYVLMCDGVDLQKQVEDFVSAQTWHQQFKTIGVMLFQQDKQFIYPLI
 25 HIPKIDSLIWENSCGSGAASIGVLVNYLTDHDIQDYLNVNQPGGSIIVSSRKSGQNEYQTT
 IKQVSTVATGQAYIEQETMTQI*

Sequence 2751

Contig_0813_pos_3664_2447,

30 putative peptide of unknown function

atggttggtagtggaacgggtcgctattcaacttgctcgactatgtcatttacatggagaa
 catatagttgatatggtgagtcgcttcatgcatcaaccaaataagagagtcctttgat
 gcttatcaacgtgacggctttttttcagtaatgactcaaaatgatgcacatcagtgtttt
 tcaggtaagtttacggtagacatttttttaagatgttaagatattactgaatattat
 35 gacgtggtgatttttagcatgtactgccgatgcgtatcgaccgatattacagcaattatct
 aagtcacattaaagcgtatttaagcaaatcatcttggctccaccaacattaggtacacat
 atgcttgttaagcaattactatcagatgttcaatgtgaagtggaagtgttccattttcc
 acttatctaggcgatacccgaaatatttgataaagcacaaccaattgtgtcctaaacaca
 cgagttaaatcaaaattattcgtaggttcgactcaatctcagtcctatgacgttgtgtaag
 40 cttaagtctttatttgactatttgaatatagaattaacaacgatggacacaccactacat
 gcggagatacataatagttcactttatgtacaccaccattgtttatgaatcaattttca
 ttaaaggcggattttgaagggaacgaaagtaccagtatatgtatataagctattttccagag
 ggtccaatcacaaatgaccttaatacacgaaatgcgattaatgtggcaagaaatgatgatg
 atattaaaaaaattaaaggtaaccttcggtaactcttaaaagtttatggtgaaagaaaac
 45 taccctatacgttatgagaccatgcgcgaagtagatattgaaaactttaaaaatttacca
 gctattcatcaagagtatctactttatgtgcgatatagcaatttttaacgatccgttt
 tctaataccggacgatcaaggtgcataattttgattttctgcggtaccatacaaacatgtt
 gatactgatgaacaaggagtcatacatataccacgcgatgccgagtgaagattattatcgt
 actttgataaattcaagcgatttgaagagcattaaacggttgcaacaccgatgattgacaca
 50 ttgttattacgttatgaaaatactgttaacaataactgtgacacacatttacatcaacaa
 ctatattccctaaattga

Sequence 2752

MVSGSPVAIQARLCHLHGEHIVDMVSRVHASTKSKRVFDAYQRDGFSSVMTQNDHAHQCF
 55 SGKFTVRHFFKDVKDITEYYDVVILACTADAYRPILQQLSKSTLKRKQIILVSPTLGS
 MLVKQLLSDVQCEGEVISFSTYLGDTIRFDKAQPHCVLTVRKSKLFFVGSTQSQSMFLCK
 LKSGFDYINIELTMDTPLHAEIHNSSLYVHPPLFMNQFSLKAVFEGTKVPVYVYILFPE
 GPITMTLIHEMRMLWQEMMILKKLVPSVNLKFMVKENYPIRYETMREVDIENKNLP
 AIHQEYLLVRYTAILIDPFSNPDDQGAYFDFSAVPYKHVDTDEQGVIIHPRMPSEDIYR

TLIIQAIGRALNVATPMIDTLLRLRYENTVKQYCDTHLHQQLYSLN*

Sequence 2753

Contig_0814_pos_338_709,

5 is similar to (with p-value 8.0e-51)
 >gp:gp|D76414|D76414_4 Staphylococcus aureus gene for histidyl-tRNA synthetase, ppGpp hydrolase, lytic enzyme, complete cds. NID: g2580431.
 10 atgactcgttccaatgataaatatgtgtccctagatgatagaaatataaagggtgacgca
 tttattagtatacataatgatgcgcttgattcttcaaagccaatggagtgactgttat
 15 tgggttaagacaagcaagaatcacttgcacaaactttaattctgcaattcaaaagaag
 gcattattgacaaatcgaggttctagacaacaaattatcaagtgttgagacagacagat
 ataccagcagtagctgttagaggttaggctatataagtaatcctactgatgaatcaatgatt
 aatgatcaattacatagacaagtggttgaacaagctattgttgatgggtttaaaacaatat
 ttctcgtcctag

Sequence 2754

MTRSNDKYVSLDDRNKGDFAFISIHNDALDSSNANGVTVYWFKDKQESLAQTLNSAIQKK
 20 ALLTNRGSRQQNYQVLRQTDIPAVLLELGYISNPTDESMINDQLHRQVVEQAIVDGLKQY
 FSS*

Sequence 2755

Contig_0814_pos_1081_2355,

25 is similar to (with p-value 0.0e+00)
 >sp:sp|O32422|SYH_STAAU HISTIDYL-TRNA SYNTHETASE (EC 6.1.1.2
 1) (HISTIDINE--TRNA LIGASE) (HISRS). >gp:gp|D76414|D76414_5
 Staphylococcus aureus gene for histidyl-tRNA synthetase, ppG
 pp hydrolase, lytic enzyme, complete cds. NID: g2580431.
 30 atgattagatgccaagaggtacccaagatatcttgccgcaagattctgctaaatggcgt
 tacattgaaaatcgattacacacattaatggaattgtataattataaagaaataagaacg
 ccaatttttgaaagtactgaactttttgcaagaggcgtgggggattctactgacgttggt
 caaaaggaaatgtatacatttaaagataaaagggatcgtagttaacattacgtcctgaa
 ggaactgcagccgttgtagcttcatatattgaacacaaaatgcaaggtgaaccaatcaa
 cctatcaaaactttactacaatggctcctatgtttagatacgaacgtaaacaaaaaggaaga
 35 tatcgccaatttaaccaatttgggtgtcgaagcaataggagcggaaaaatcctagtattgat
 gctgaaataactcgctatgggttatgcatatataatgagtccttcggattaaagcatttaaag
 ttagttatcaaatagatttggtgatagtgaaatcacgtaaaagaatataacgaagcattagta
 aaacattttgaacctgtgattgatacattttgttcagattgtcaatcaagattacacact
 aatcctatgagaatttttagattgtaaaatcgatagagataaagaagcagtaaaaaatgca
 40 ccccgtagcacagattatctcaataatgattctaaatcctattatgaacaagttaaatta
 catcttgataatttgaacatatcttatgttgaagatcctaacttagttcgtgggttagat
 tattatactcatactgcctttgaattaatgattgataatccagagtagatggagctatc
 actacattatgtggtggtggtcgatataatgggttggttacaattattagatgggtccagat
 gaaacaggtattgggtttgcactaagattgaaagattattgatggcacttgatgaagaa
 45 ggtattttcattagatgtaagtgaagattttgatttattttgttgacaatgggagaagat
 gccgatcggttatgctgttaagttaataatgatttaagaagaatggaataaaagtagat
 aaggattatctaacaagaaaaattaaaggacaaatgaaacaagctgaccgtcttaatgct
 aaatatacagtagtaattggagatcaagagcttgaaaataatgaaattggtgtgaaaaac
 atgacattcaggcgaatcagaaaaatgtacaattagacgaattgggttaattattttaaaagt
 50 agaaaaggaagtctaa

Sequence 2756

MIKMPRGTDILPQDSAKWRYIENRLHLMELYNKYKEIRTPIFESTELFARGVGDSTDV
 55 QKEMYTFKDKGDRSLTLRPEGTAAVVRSYIEHKMQGEPNQPIKLYYNGPMFRYERKQKGR
 YRQFNQFGVEAIGAENPSIDAEILAMVMHIYESFGLKHLKLIVNSIGDSESRKEYNEALV
 KHFEFVIDTFCSDCQSRHLTNPMRILDCKIDRDKAVKNAPRITDYLNNDSKSYEQVKL
 HLDNLNLSYVEDPNLVRGLDYTHAFELMIDNPEYDGAITTLCGGGRYNGLQLLDGPD
 ETGIGFALSIERLLMALDEEGISLDVSEDFDLFVVTMGEDADRYAVKLINDLRRNGIKVD
 KDYLNRKIKGQMKQADRLNAKYTVVIGDQELNNEIGVKNMISGESENVQLDELVNYFKS

RKEV¹

Sequence 2757

Contig_0815_pos_2348_3526,

- 5 is similar to (with p-value 1.0e-92)
 >sp:sp|P38494|RS1H_BACSU_30S_RIBOSOMAL_PROTEIN_S1_HOMOLOG. >
 gp:gp|U11687|BSU11687_5 Bacillus subtilis 168 jofA, jofB, Ms
 sA homolog (jofC) and ribosomal protein S1 homolog (jofD) ge
 nes, complete cds, and joeB gene, partial cds. NID: g533101.
 10 >gp:gp|Z99115|BSUB0012_228 Bacillus subtilis complete genom
 e (section 12 of 21): from 2195541 to 2409220. NID: g2634478
 . >gp:gp|L47648|BACSERA_20 Bacillus subtilis phosphoglycerat
 e dehydrogenase (serA), ypaA, ferredoxin (fer), ypbB, recS,
 ypbD, ypbE, ypbF, ypbG, ypbH, glutamate dehydrogenase (ypcA)
 15 , ypdA, ypdB, ypdC, spore cortex lytic enzyme (sleB), ypeB,
 ypfA, ypfB, cytidine monophosphate kinase (cmk), ypfD, ypgA,
 yphA, yphB, yphC, NAD+ dependent glycerol-3-phosphate dehyd
 rogenase (glyc), yphE and yphF genes, complete cds. NID: g11
 46195. >gp:gp|L47648|BACSERA_20 Bacillus subtilis phosphogly
 cerate dehydrogenase (serA), ypaA, ferredoxin (fer), yphB, r
 20 ecS, ypbD, ypbE, ypbF, ypbG, ypbH, glutamate dehydrogenase (
 ypcA), ypdA, ypdB, ypdC, spore cortex lytic enzyme (sleB), y
 peB, ypfA, ypfB, cytidine monophosphate kinase (cmk), ypfD,
 ypgA, yphA, yphB, yphC, NAD+ dependent glycerol-3-phosphate
 25 dehydrogenase (glyc), yphE and yphF genes, complete cds. NID
 : g1146195.
 atgactgaagaattcaatgaatcaatgattaatgatattaaagaaggtgacaaagtcact
 gttgaagttcaacaagtagaggataaacaagttgttgtgcatattaatggtggcaaat
 aatggaattattcctatttagccagctttcaacacatcatatcgaaaaccctagtgaagtt
 30 gtaaaagtcggtgatgaagtcgaagcatatgtcactaaaatcgagttcgacgaagaaaat
 gatactggggcatacatattttatcaaaaagacaacttgaaactgaaaaatcttatgaatat
 ttacaagaaaaactagataacgatgaagtgattgaagctgaagttactgaagtagttaa
 ggtggttttagtcggtgacgttggtcaaagaggggttgacctgcttctctaatttcaact
 gatttcattgaagatttttctgtattcgatggtcaaacaatccgtattaaagtggaagaa
 35 cttgatcctgaaaacaatagagtcatttttaagccgtaaaagctgtggaacagttagaaaac
 gacgctaataaaagcttcaatattagattctttaaataagggcgatgttattgatggtaaa
 gttgctcgattaactaaacttttggtgctttcattgatattggtggcgtagatggtttagtt
 cacgtttctgaattatctcatgaacatgttcaaaccaccagaagaagttgtgtcagtaggt
 gaagcagtc aaagttaaagttaaatctgtagaaaagattctgaacgtatttctttatct
 40 attaaagacactttaccaacaccatttgaaaacattaaagggaaatttcacgaagatgat
 gttattgaaggtactgtagtacgtttggcgaaactttggcgcatctcgtagaaattgtcca
 tccgtccaaggttttagtgcataatttctgaaatcgatcataaacatatcggttctcctaac
 gaagtattagaaccttgacaacaagttaatgtaaaaatattaggtatcgatgaagataat
 gaaagaatttccattatcaatcaaagcaacgttacctaaagaaaatgtcattgaaagtac
 45 gcattccacaactcaatcatatcttgaagatgataatgatgaagataaaccaacattaggc
 gatgtttttggtgataaatttaaagaccttaagttttaa

Sequence 2758

- 50 MTEEFNESMINDIKEGDKVTVEVQVEDKQVVVHNGGKFNIGIIPISQLSTHHIENPSEV
 VKVGDEVEAYVTKIEFDEENDTGAYILSKRQLETEKSYEYLQEKLDNDEVIEAEVTEVVK
 GGLVVDVGQRGFVPASLISTDFIEDFSVFDGQTIRIKVEELDPENNRVILSRKAVEQLEN
 DAKKASILDLSNEGVDIDGKVARLTNFGAFIDIGGVDGLVHVSELSHEHVQTPEEVVSVG
 EAVKVKVKSVEKDSERISLSIKDTLPTPFENIKGFHEDDVIEGTVVRLANFGAFVEIAP
 SVQGLVHISEIDHKHIGSPNEVLEPGQQVNVKILGIDEDNERISLSIKATLPKENVIESD
 55 ASTTQSYLEDDNDEDKPTLGDVFGDKFKDLKF*

Sequence 2759

Contig_0815_pos_4093_5052,

is similar to (with p-value 0.0e+00)

>sp:sp|P50743|YPHC_BACSU HYPOTHETICAL 48.8 KD GTP-BINDING PROTEIN IN CMK-GPSA INTERGENIC REGION. >gp:gp|Z99115|BSUB0012_224 Bacillus subtilis complete genome (section 12 of 21): from 2195541 to 2409220. NID: g2634478. >gp:gp|L47648|BACSERA_24 Bacillus subtilis phosphoglycerate dehydrogenase (serA), ypaA, ferredoxin (fer), ypbB, recS, ypbD, ypbE, ypbF, ypbG, ypbH, glutamate dehydrogenase (ypcA), ypdA, ypdB, ypdC, spore cortex lytic enzyme (sleB), ypeB, ypfA, ypfB, cytidine monophosphate kinase (cmk), ypfD, ypgA, yphA, yphB, yphC, NAD+ dependent glycerol-3-phosphate dehydrogenase (glyc), yphE and yphF genes, complete cds. NID: g1146195. >gp:gp|L47648|BACSERA_24 Bacillus subtilis phosphoglycerate dehydrogenase (serA), ypaA, ferredoxin (fer), ypbB, recS, ypbD, ypbE, ypbF, ypbG, ypbH, glutamate dehydrogenase (ypcA), ypdA, ypdB, ypdC, spore cortex lytic enzyme (sleB), ypeB, ypfA, ypfB, cytidine monophosphate kinase (cmk), ypfD, ypgA, yphA, yphB, yphC, NAD+ dependent glycerol-3-phosphate dehydrogenase (glyc), yphE and yphF genes, complete cds. NID: g1146195.

gtgaataaagttgataatcttgaaatgcgtaatgatctatgatttctattcttttaggc
 tttggagatccatctctatttctggttcacatggattaggacttgagatttgctagat
 gcagttgttgaaaactttaataagaatcagaagatccttatgacgaagatacgatacgt
 ctttctatcatcggtagacctaatgtttgtaaatctagcttggtcaatgctatttttaggc
 gaagaacgtgttattgtgtctaattgttgcgtggtacaactcgagatgccattgataccgag
 tactcttatgatggacaagattatgtattgattgatactgctggaatgagaaaaaagggt
 aagggtgatgaatcgactgaaaaatattctgtattacgtgcattaaaagcgattgagcgt
 tcagaagtagtatttagtattatcgatgctgaacaaggtataattgaacaagataaacgt
 gtagctggctatgcacatgaggaaggtaaagctattgtcattgtagtaataaatgggat
 acagttgaaaaagatagtaagacaatgaaaaaattcactgatgatgtagaaatgaattt
 caatttttagattatgctcaaatcgcggttcgtatcagcaaaaagaagggttaagattaaaa
 acattattcccttatatacaatcaagccagtgaaaatcataaaaagcgtgtccaaagttct
 acactaaatgaagttgttactgatgccatctctatgaatccaacactactgacaaaagggt
 agaaatgaatgtattctatacaactcaggttgcaattgaaccaccgacatttgtagta
 tttgtcaatgatgttgaaatgaatgcatttttcttataggagatatttagaaaaatcaaata
 cgtaatgcttttggttttgaaggaacacctattcatattattccaagaaaaagaaattaa

Sequence 2760

VNKVDNLEMRNDIYDFYSLGFGDPYPISGSHGLGLGDLDDAVVENFNKESEDYPYDEDTIR
 LSIIGRPNVGKSSLVNAILGEERVIVSNVAGTTRDAIDTEYSYDGGQDYVLIDTAGMRKKG
 KVEYESTEKYSVLRALKAIERSEVVLVVIDAEQGIIEQDKRVAGYAHEEGKAIVIVVNKWD
 TVEKDSKTMKKFTDDVRNEFQFLDYAQIAFVSAKEGLRLKTLFPYINQASENHKKRVQSS
 TLNEVVTDIAISMNPTPTDKGRRNLNVFYTTQVAIEPPTFVVFNDELHMFYSYRRYLENQI
 RNAFGFEGTPIHIIPRKRN*

Sequence 2761

Contig_0815_pos_5073_6071,
 is similar to (with p-value 0.0e+00)
 >gp:gp|Z99115|BSUB0012_223 Bacillus subtilis complete genome
 (section 12 of 21): from 2195541 to 2409220. NID: g2634478.
 >gp:gp|L47648|BACSERA_25 Bacillus subtilis phosphoglycerate
 dehydrogenase (serA), ypaA, ferredoxin (fer), ypbB, recS, ypbD,
 ypbE, ypbF, ypbG, ypbH, glutamate dehydrogenase (ypcA),
 ypdA, ypdB, ypdC, spore cortex lytic enzyme (sleB), ypeB, ypfA,
 ypfB, cytidine monophosphate kinase (cmk), ypfD, ypgA, yphA, yphB,
 yphC, NAD+ dependent glycerol-3-phosphate dehydrogenase (glyc),
 yphE and yphF genes, complete cds. NID: g1146195. >gp:gp|L47648|BACSERA_25 Bacillus subtilis phosphoglycerate dehydrogenase (serA), ypaA, ferredoxin (fer), ypbB, recS, ypbD, ypbE, ypbF, ypbG, ypbH, glutamate dehydrogenase (ypcA), ypdA, ypdB, ypdC, spore cortex lytic enzyme (sleB), ypeB, ypfA, ypfB, cytidine monophosphate kinase (cmk), ypfD, ypgA, yphA, yphB, yphC, NAD+ dependent glycerol-3-phosphate dehydrogenase (glyc), yphE and yphF genes, complete cds. NID: g1146195.

pcA), ypdA, ypdB, ypdC, spore cortex lytic enzyme (sleB), ypeB, ypfA, ypfB, cytidine monophosphate kinase (cmk), ypfD, ypgA, yphA, yphB, yphC, NAD⁺ dependent glycerol-3-phosphate dehydrogenase (glyc), yphE and yphF genes, complete cds. NID: g1146195.

5 atgagaaaaattacagtttttggtagttttggtagtgcattagctaattgtatta
gctcaaaatggtagtcatgatttttaattgtggggcaaaaatgtagagaatgtagatgaactt
aacacacatcatatgaacaaaaattatcttaagatgctaaattagattcatctataaaa
gcaactgtagatttaataaaggcagtgcaattttcagatatctatcttatggcactacct
10 acaaaagcaattagagaagtatcaaaagatatcgatcaattactcacatctaaaaagact
tttattcatgttgctaaaggcattgaaaacgatacatttaagcgctatctgaaatgatt
gaggactctatctcttcagaacataatggaggaatcggcgtcttatcaggtccaagtcatt
gctgaagaagttgttataaaacaacctacaactgtagctgcatcatctaaagataataat
gtgagcaaaacttattcaagatttatttatgaacgactatttacgtgtttacacaaataat
15 gatttagtaggtgtagaatttagtggtgctttaaaaatattatagctatagctagtggt
atcgttgccgcatgggttacggtgataaatgcaaaagcagctttaatgacacgaggttta
gccgaaatcagtcgacttggtgagaaaacttggtgcagatccaatgactttcttaggtcta
ggtggcataggtgacttgatcgtaacttgtagctccacacattcacgaaattacacactt
gggttttaattagggcaaggcaaacagcagaagaagctttaaaagagatgaaaatggtg
20 gttgaaggtatttatacaactaaatcagtatatcatcttgctcaacaagaaggagtagag
atgcctatcactaacgcatttatgaagttttatttgaagatgtccctgtaagtaaaagt
gttagaacacttatggaaagagacaaaaaagcagaataa

Sequence 2762

25 MRKITVFGMGSFGTALANVLAQNGHDVLMWGKNVENVDELNTHHMKNYLKDAKLDSSIK
ATVDLNAKAVQFSDIYLMALPTKAIREVSKDIDQLLTSKKTFIHVAKGIENDTFKRVSEMI
EDSISSEHNGGIGVLGSPSHAEVVIKQPTTVAASSKDNVSKLIQDLFMNDYLRVYTNN
DLVGVELGGALKNI IAIASGIVAGMGYGDNAKAAALMTRGLAEISRLGEKLGADPMTFLGL
GGIGDLIVTCTSTHSRNYTLGFKLGQGKTAEALKEKMMVVEGIYTTKSVYHLAQQEGVE
30 MPITNALYEVLFEFVPSKSVRTLMEKDKAE*

Sequence 2763

Contig_0815_pos_0_368,
is similar to (with p-value 2.0e-21)
35 >sp:sp|P42086|PBUX_BACSU_XANTHINE_PERMEASE. >pir:pir|S51310|
S51310 xanthine permease - Bacillus subtilis >gp:gp|L77246|B
ACYACA_3 Bacillus subtilis (YAC10-9 clone) DNA region between
the serA and kdg loci. NID: g1256615. >gp:gp|Z99115|BSUB00
12_147 Bacillus subtilis complete genome (section 12 of 21):
40 from 2195541 to 2409220. NID: g2634478. >gp:gp|X83878|BSXPT
PBUX_2 B.subtilis xpt and pbuX genes. NID: g633168.
atgtgcgggtagcgacatttcttcaagcaataaagtcacagggactggattaccgatt
gtactaggatgtacgtttactgcccgttgacacctatgatactcatcggtcaaacgaaaggga
cttgatgttttatatggttcgcttttaatatccggtatcttagttgttttaattgcacct
45 tttttctcttatttagttaaatctttccacctgttgtaacaggaagtgttgtagacaatt
attggaatcaatttaattgacagttgcaatgaattacttggcaggtggtgaaggagcgaaa
aactatggcgatactaagaatttaattaggtggtgttacactactcattattcttatt
TGATTTAT

Sequence 2764

50 MCGVATFLQANKVTGTGLPIVLGCTFTAVAPMILIGQTKGLDVLVYGSLLISGILVVLIAPI
FFSYLVKFFPPVVTGSSVVTIIGINLMPVAMNYLAGGEGAKNYGDTKNLILGGVTLIIILI
*FX

Sequence 2765

55 Contig_0818_pos_4072_3692,
putative peptide of unknown function
gtgtttattttgttaacatttggattttatgtattttttgctggccataataatccaggt
ggtggctttattggtggcttgatttttagctcggcatttatcttaattgtttcttgccctt

gatgtaaatgaagtgttgaaaagccttgccattgattttaaaaaattaatgattataggt
 tcactcatatctgttgcaactgcatcagtcctatgttttttgggaagccatttttatat
 caaactgaagcaaatgtaacatttccattactaggacatgttcatgttactactgtgact
 ttatttgagcttgccatcttattaacagtagtaggtgtgattgttacagttatgctatct
 5 ataagtggggttagatcatga

Sequence 2766

VFILLTFGFYVFFAGHNNPGGGFIGGLIFSSAFILMFLAFDVNEVLKSLPIDFKKLMIIG
 SLISVATASVPMFFGKPFYQTEANVTFPLLGHVHVTTVTLFELGILLTVVGVIVTVMLS
 10 ISGGRS*

Sequence 2767

Contig_0818_pos_3674_3351,
 is similar to (with p-value 3.0e-20)
 15 >gp:gp|AB015981|AB015981_4 Staphylococcus aureus genes for O
 rfa, MnhA, MnhB, MnhC, MnhD, MnhE, MnhF and MnhG, complete c
 ds. NID: g4001723.
 gtgataggatttttagtgatttattggaacttatatgattttatctattaatttaattcgt
 attgttattggtattttctatttatacacacgcccgaatttaattattatgagtatggg
 20 aaatatggacctcatatgtctgaaccgctaattcaaggctcatgctcaaaactttgttgat
 ctttattacaagctatcgttttaacagctattgtgattggatttggtatgactgcgttt
 ttattggtgttaatatagaaacttacagagtaactaaaggagatgaaataagtgcattg
 aaaggtgatgaagatgatgagtaa

Sequence 2768

VIGFLVF⁷GTYMILSINLIRIVIGISYTHAGNLIIMSMGKYGPHMSEPLIQGHAQNFVD
 PLLQAIVLTAIVIGFGMTAFLLVLIYRTRYVTKEDEISALKGDEDE*

Sequence 2769

Contig_0818_pos_3334_1859,
 is similar to (with p-value 2.0e-62)
 30 >gp:gp|AB015981|AB015981_5 Staphylococcus aureus genes for O
 rfa, MnhA, MnhB, MnhC, MnhD, MnhE, MnhF and MnhG, complete c
 ds. NID: g4001723.
 atgttggtgccttttgatgtgcttttaatttttagtcttcactaaaaataaaaaatcgatt
 tcgaaaatcctatccattacaactatgattgttaatacaatgatttcaattgctttactt
 atttatgtcggttaatacataaaccgataaacttgattttgggggatggaaagcacctttc
 35 ggcattcaatttctaggtgattcactgagtcgtctgcttatgggtgcagtatcatctttgtt
 gttacgctaataatggcatacggctttggttagaggggagaagcgagtcattcgatttcac
 40 ctccctacatttattctttttaacagtaggtgttattggttcggttttaacttctgat
 ttatttaacctatacgtgatgtttgaaattatgcttcttgcgttgacttggttaca
 ttaggacaatctgttgaacaattacgtgcagcgatagatatgttggttctgaatatttta
 ggttcgtgggttgcttttattaggaattggcatgttatataagacagtcggaacactta
 ttctcacatttagcgatgcgattgaatcatatggaaaataaccaaacaataacgatgata
 45 tcttttagtatttctagttgcttttagttcaaaggcagcactagtgattttcatgtggtta
 cctaaagcatatgcagtgcttaatacggaaacttgccgcgttatttgcagcattgatgaca
 aaagttggagcttatgcgcttattcggtttttactttactattcgaccatcatccaagc
 gtcacgcatacattgctcggtttatggcttgatcacatgattatcggtgcatttggt
 gtcacgcttacaaagatattaagaaaattgcggcttatcaagttattttgtctattgga
 50 ttcatatttttaggttttaggttctcatactatatacagggtgtaaatgggtgctatcttctat
 tttagcgaatgatattatcggttaagacattattgttttttgtaattggtagtctgtttat
 atgtcaggctatcgaaattatcagttttaagtggactggcaaaaagagaaccattcttt
 ggtgttgcatttgtcgtggtaattttgtctataggtggcgtacctccttttagtggttt
 ccgggttaaagtccttaattccaaggggctattacaaatggtaattatattggttttagca
 55 cttatgattgtgacaagtttaattgctatgtatagtccttttagagtgatgtttataatg
 tattttgggtgatgctgacggagaacaagtacaatttagaccactacctatttatcgtaaa
 ggtttacttagtgttttagttgtagtggttattagcgatgggtattgcagccctgttgtt
 ctgaaagtaacagaggatgcaacaattcttaatatgaaagaagatgtctttcaaaagaat
 gtaaatcacatttgaaggagggttaatacataagtga

Sequence 2770

MLLPFVCALILVFTKNKNRISKILSITTMIVNTMISIALLIYVVNHKPITLDFGGWKAPF
GIQFLGDSLSLLMVSFVVTLMAYGFRGEKRVNRFHLPTFILLTVGVIGSFLTSD
5 LFNLYVMFEIMLLASFVLVTLGQSVEQLRAAIVYVVLNILGSWLLLLGIGMLYKTVGTLN
FSHLAMRLNHNMQTITMISLVFLVAFSSKAALVIFMWLPKAYAVLNTELAALFAALMT
KVGAYALIRFFTLFDHHPVTHTLVFMACITMIIGAFGVIAVKDIKKIAAYQVILSIG
FIILGLGSHSISGVNGAIFYLANDIIVKTLFFVIGSLVYMSGYRNYQYLSGLAKREPFF
GVAFVVVIFAIGGVPPFSGFPGKVLIFQGAITNGNYIGLALMIVTSLIAMYSLFRVMFIM
10 YFGDADGEQVQFRPLPIYRKGLLSVLVVVVLAMGIAAPVVLKVTEDATNLNMKEDVFQKN
VNTHLKEVNHK*

Sequence 2771

Contig_0818_pos_1751_1380,
putative peptide of unknown function
15 gtgatttatattctgcatcgcttttttggtgaagaattttatttgaaaaagatatgggtg
gctattaaatttttagctgtatacctataccagcttattacttctagtataagtaaccata
aattacatcttatttaagacgaatgaagtttaattccaggtttactcacatatgaaacttca
ttaaaaagtaattgggctattactttttaacgattttaattattattactccaggatcg
20 acagttatcgaatttctaaaaataactaataaattttttatttcacagttattgatgtgtca
gaaaaagataaagaaaatcttctaaaaagtatttaagcagttatgaggatttaattttggag
gtgacacgatga

Sequence 2772

25 VIYILHRFFGEEFYLKKIWVAIKFLAVYLYQLITSSISTINYILFKTNEVNPGLLTYETS
LKSNEWAITFLTILIIITPGSTVIRISKNTNKFHISIDVSEKDKENLLKSIKQYEDLILE
VTR*

Sequence 2773

30 Contig_0819_pos_5530_6732,
is similar to (with p-value 0.0e+00)
>sp:sp|Q53634|MENE_STAAU O-SUCCINYLBENZOIC ACID--COA LIGASE
(EC 6.2.1.26) (OSB-COA SYNTHETASE) (O-SUCCINYLBENZOATE-COA S
YNTHASE). >gp:gp|U51132|SAU51132_1 Staphylococcus aureus o-s
35 uccinylbenzoic acid CoA ligase (mene), and o-succinylbenzoic
acid synthetase (menc) genes, complete cds. NID: gl255258.
atgataaatacacggtttaacgcgacatgagatgataaatcaaatgaattcagtcgacata
gcaacgattgtacacacggttgcttttagaattagaagggtttaatttatatcattttaat
gatttaacacaattagataaacatgatgtttcaggttacaatttaatttagaatcgatt
40 gcatcaattatgtttacgtctggaacgacgggacctcaaaaagctgtgctcaaaagttt
aataatcatttagccagtgctaaaggctgtaaacaaagtttaggattcgaacaaaatact
gtgtggctttcggctcttacctatatatcatatttctgggctcagtggtattttgcgcgca
gtgatagaaggattcactgtcagacttggttaaaaagtttcaaactgatgatatgttaaca
caaataaagacttatccaatcacccatattgctccttggtccacaaacgttaaagtgggta
45 atggatgcaggattgactcaaccattttcttttagaaaaaattctgctaggtgggtgctaaa
ttatcaccacaatttaattgagcaagcattgacttatcggtttacctgtatataattctttt
ggatgatgacagaaacttgctctcagtttctaacagcctcacctcaaatgctcaaagaacgt
ttcgatactgttggaaaaccaagtgaatatgtcgaagtgaataaaaaatcccaacgca
tatggacattggagaggttatttaattaaagggtgaatatgtgatgaatgggtatttatatccc
50 aaattatataaaagacacatttgataatgatgggtattttcaaactggagatatagctgaa
atagatgatgaagggtacgtcataatatatgatcggcgcaagatttgattataagtggt
ggagagaatatttatccttaccaaattgaaacaatcgaaaagactttgaaggcattgaa
gatgccgtatgtgtaggaatatcagatgatacttgggggtcaagtaccaatattatattat
gtgacaaatcaagatatataatcaaactgaattaatagaacattttgagaatcatttagct
55 agatataaaaattcctaaaaaatattatcaggtcaaatctttaccttatatcgcagcaggt
aaattacaacgtaaaaagggtcaaaagtgaagacttgatgagggaaagaataatgaaagt
taa

Sequence 2774

MINTRLTRHEMINQMNSVDIATIVHTLPLELEGFNLYHFNDLTQLDKHDVSGYKFNLESI
 ASIMFTSGTTGPQKAVPQTFNNHLASAKGCKQSLGFEQNTVWLSVLPIYHISGLSVILRA
 VIEGFTVRLVKKFQTDMLTQIKTYPITHMSLVPQTLKWLMDAGLTQPFSLKILLGGAK
 LSPQLIEQALTYRLPVYNSFGMTETCSQFLTASPQMLKERFDTV GKPSENVKIKPNNA
 5 YGHGELLIKGENVMNGYLYPKYLKDTFDNDGYFQTGDIAEIDDEGYVIIYDRRKDLIISG
 GENIYPYQIETIAKDFEGIEDAVCVGISDDTWGQVPILYYVTNQDINQTELEHFENHLA
 RYKIPKKYYQVKSLPYTSTGKLQRKKVKSEDLNEGKNNES*

Sequence 2775

10 Contig_0819_pos_6791_0,
 is similar to (with p-value 6.0e-43)
 >gp:gp|U51132|SAU51132_2 Staphylococcus aureus o-succinylben-
 zoic acid CoA ligase (mene), and o-succinylbenzoic acid synt-
 hetase (menc) genes, complete cds. NID: g1255258.
 15 gtgaagttaacgcacatcgagaaagtttgtttacggaaatagtaacttatagtggagaaact
 tattatggagaatgtaatgcatttttaactaattggatgataaagaacaataactcaca
 gtcgtaaacagattaagacagtgataaccgcaagtacttcataaagatatgacatctttt
 gattcatggttacctaataatcaaatgaatgatgcgccagctgctagatcaatgggt
 gtcattggctgtttatcaaatgtataacgacttgcattgattttgaagtacaatacgggtgct
 20 acagttagtggtttaaactaatagtcgaattgaaacattattagaaacaagaccgaaacgt
 ataaaacttaaatggtaacatcactcatcaaatgctgaaactatacgtttattaaat
 tttgattgtgatattgctatagatgcaaatgaatcattaacaaagccatcatttttaca
 ttacccaacgtaatacatcagatattatatattgaagaaccttttaaaattct

Sequence 2776

25 VKLTHRESLFTEIVTYSGETYYGECNAFLTNWYDKETILT VVNRLRQWIPQVLHKDMTSF
 DSWLPYLNQMNDA PAARSMVVMVAVYQMYNDLHDFEVQYGATVSGLTNSQIETLLETRPKR
 IKLKWSTSLIKDLETIRLLNFDCDIAIDANESLTKPSFLQLANVNTSDIYIEEPFKIL

Sequence 2777

30 Contig_0819_pos_4059_3394,
 putative peptide of unknown function
 atgcaacaagaaacgacatcatggtacaaacaagaatggtttatagttttatcactttta
 ttcatttttccactaggtttattttctcatgtggaaatttagcaagtggccatctattgca
 35 agaacaatcattactgttgcaattttcagttatcgtatttagcaagcattacctattatggt
 aatctacaaatgattgtaccagcaacatcaaattcaataacgaaactaaagaaactaca
 gagaataatgtaaatgataaagacgagcgaatcataaaactgcagtagaagaaacaaa
 actaattatgactccaccaaaagaaaataactaaagaacctggaaaagaaatgaatctgca
 acacgattggagaactctgcgcttgaaaaggcaaatcatattatgatgattttcacatg
 40 tctaaactaggaatttatgatattttaacatctgaatatggagaaaaatttgataaagaa
 gatgcacaatatgctatagatcatctagaggctgattatgaaaagaatgcacttgagaaa
 gcaaaatcatatgccaaagatatgcataatgtctaatactcaatttacgatcttttggtg
 tctaactacgggtgaaaaatttacagaatcagaagcaaaatatgctattgagcatttgga
 aattaa

Sequence 2778

45 MQQETTSWYKQEFIVLSLLFIFPLGLFLMWKFSKWPSIARTIITVAISVIVLASITYYG
 NLQMIVPATSNNSNETKETENNNDKDERNHKTAVEETKTNYDSTKENTKEPGKENESA
 TRLENSALEKAKSYDDFHMSKLGIDILTSEYGEKFKEDAQYAIHLEADYEKNALEK
 50 AKSYAKDMHMSNDISIYDLLVSNGEKFTESEAKYAIEHLDN*

Sequence 2779

Contig_0819_pos_2417_1683,
 putative peptide of unknown function
 55 atgagaaaattattttctatcaattttatctattataataatttcaagttttgtgttgca
 acagggtttcaaaacgttaattgctgcaataatgaggtgagcaaacctcaaagtaattgtg
 gatagtaaaactaaacaaatattattaaaaaataaaaaaatctaattgcttataaaaag
 cagctaattcaacttctattgactcaataaaagatgacgatattattgttcatttgga
 aaaggtaaaaacaccaatgtatattctattaactatgtatttggtaaaaaattagctaaa

atgaatgataaccttgcaatgattgaattaaaatataacgaagcaaacaatgaagttttc
 tacaatcatatgatgtattcaaaatttgttaaatataataaagaatatattaatgatg
 aaaggcattttaaatggcaaaccttattatgaatttaatatagatcagaaaggctc:ttac
 tacgataaaaacttttaacataacctctaaagatgaaatcgaaaaagattctgctaaaaac
 5 ttaccacctaagagcggtggttggtgtgaatgggcagtaggagctttatgtggtaccggt
 ggagctgcaggttgttgggcaactgctacagctttaggtattactactggttggggagggc
 ttttcattagctacaatttgtggtctgataagctctctaggtgtactggtgcaaccaac
 tatatttgtaataa

10 Sequence 2780

MRKLFLSILSIIIISSFCVATGFQNVNAANNEVSKPQSNVDSKTKQNIKKIKKSNAKK
 HANSTSIDSIKDDDIIVHLDKGKNTNVYSINYVFGKLLAKMNDNLAMIELKYNEANNEVF
 YNHMMYSKFVKYNNKEYINMKGILNGKPYEFNIDQKGHYDKNFKHTSKDEIEKDSAKN
 LPPKERGWCEWAVGALCGTGGAAGCWATATGITTGWGGFSLATICGLISSLGCTGATN
 15 YICK*

Sequence 2781

Contig_0822_pos_6711_5857,
 is similar to (with p-value 2.0e-42)
 20 >gp:gp|D78193|BACGNTZA_33 Bacillus subtilis 36kb sequence be
 tween gntZ and trnY genes encoding 34 ORFs. NID: g1064780. >
 gp:gp|Z99124|BSUB0021_145 Bacillus subtilis complete genome
 (section 21 of 21): from 3999281 to 4214814. NID: g2636442.
 atgaagtggcttaaacactacaatcccttcacacgaaactcggttattgtttatgtacta
 25 ctcatattattggtatgcaaatcatcggtttgtattttacgaatagtttagaaaaggaa
 ttactcgataactcaagaagaacataacacaatatgcgaagcaattagacgtcaatatt
 gaaaagggtttataaagataaagataaagggttcagtcaacgctcaaaaggatatccaagac
 cttttgaatgaatatgcgaatcgccaagaaataggagaaatcgctttattgataaagac
 caaattatcatggcaacaaccaagcagctcaaccgtggtcttatcaatcaaaagggttaac
 30 gacggttcagttcaaaaggcgctctccttagggcaaacgaatgatcatatggttcttaag
 gattacggaagtggtaagagcggtggttgggtatataataaccgggttaaagttgataaa
 cagacaatcggtgatatacatagaatcgaaaattaatgatgtatacaatcagctgaac
 aacattaatcagatattcatcgtagggacagcgatatcactattcattacagtaatacta
 ggattcttcattgcacgaacgattactaagccgataaccgatatgcgtaaccaaaccggt
 35 gagatgtctaaaggttaactacacgcaacgagtgagatatacggtaacgatgaaatcggt
 gagctcgcaacttgccttcaataacttatcgaaacggtgccaagaagcacaagcgaataca
 gaaagtgagaaacgtcgcttagattctgttatcacacataattacttgttgatgtatta
 cattttgattgttaa

40 Sequence 2782

MKWLKQLQSLHTKLIVIVYVLLIIIGMQIIIGLYFTNSLEKELLDNFKKNITQYAKQLDVNI
 EKVKYDKDKGSVNAQKDIQDLLNEYANRQEIGEIRFIDKDQIIMATTKQSNRGLINQKVN
 DGSVQKALSLGQTNHMLVKDYGSGKERVWVYNIPVKVDKQTIGDIYIESKINDVYNQLN
 NINQIFIVGTAISLFITVILGFFIARTITKPIITDMRNQTVEMSKGNYTQRVKIYGNDEIG
 45 ELALAFNNLSKRVQEAQANTESEKRRLDVITHNYLLNVLFHDC*

Sequence 2783

Contig_0824_pos_5618_5929,
 putative peptide of unknown function
 50 atgcacataaattcaaaattagctaaaatgccgcgtagaattgatggcaagttattctcg
 gttagtcatatcatagtaagtaaagcggttagaaaagcaaaagaagtgataggattaaac
 gataataataaactccctattcactcagacatacgcacacatcttacttactatctaaa
 ggcataccaatcgagtatataagtaaagcttttaggtcacgctactatatcacaacggtta
 gacacgtattcacatttattagaagaacataaaaaagagcaaggccaacgtgtcagagaa
 55 atattctcttga

Sequence 2784

MHINSKLAKMPRRIDGKLFVSHTSVSKAFRKAKEVIGLNDNNITPYSLRHHTSYLLSK
 GIPIEYISKRLGHATISQTLDTYSHLLEHHKKEQGQVRVREIFS*

Sequence 2785

Contig_0824_pos_4770_3757,

putative peptide of unknown function

5 atggaacgattttgtgtgttaaatacaaatcaactatattcaaatgaatccggttagaagcc
 aaattttaaaacgagcgctctaagatcatggaaaactgatcaggcagatgctcataagcct
 gcttgttttaggaccgacgcttaaacaaacagacaacttacctatacatgagttaatattc
 tttgaattaagagaacgcgctccgttttcatctagaaatcgagaatgaacaaaatcgactt
 10 aaatttcagatccttgaattactccatcaaacattccctgggttagaaagattatttagt
 agtcgatattcaatcattgcactcaacatcgagaaatctttactcatccagacatgggt
 cttgatatcgacaaggaggtactgattacacatatattcaattctacagataagggaatg
 tcaatggataaagctacaaaatatgcacttcaattaagggtgattgctcaagaaagctat
 cctaattgctgatagacattcctttctagtgcgaaaaattacgcttacttattcaacaatta
 aaacaaatcattcatcatctcaacaattagatgatgccatgattcaattagcacacaacaa
 15 ctcgattattttgaaaatattcattcgatacctggatttggttaagctaagcacagctatg
 attattggggagattgggtgatattaagcgatttaaatcaataaacaactcaatgctttt
 gttggcattgatatcaaacgatatcaatcaggtcatacacactgtagagataccatcaac
 aagcgtggtaataaaaaagcgagaaaaacttttattttgggtgattatgaatataataaga
 gggcagcatcatttatgacaatcatgtcgtcgattattactacaaactaagaaagcagcct
 20 aatgagaaaacctcataagactgccatcattgcttgataaatcgattattaaaaacaatt
 cattatcttgtaaatgaatcataaattgtacgattatcaaatgtcaccacattag

Sequence 2786

25 MERFCCVNQINYIQMNPLEAKFKTSALRSWKTDQADAHKLACLGPTLKQTDNLPHELIF
 FELRERVRFHLEIENEQNRLKFQILELLHQTFPGLERLFSSRYSIIALNIAEIFTHPDMV
 LDIDKEVLITHIFNSTDKGMSMDKATKYALQLRVIAQESYPNVDRHSFLVEKLRLLIQQL
 KQSIHHLKQLDDAMIQLAQQLDYFENIHSIPGIGKLSTAMIIGEIGDIKRFSNKQLNAF
 VGIDIKRYQSGHTRCDTINKRGNKKARKLLFWVIMNIIRGQHHDYDNDVVDYKLRKQP
 30 NEKPKHTAIIACINRLKTIHYLVMNHKLYDYQMSPH*

Sequence 2787

Contig_0824_pos_2942_2442,

is similar to (with p-value 7.0e-38)

35 >gp:gp|AF082668|AF082668_1 Streptococcus pyogenes CsrR (csrR
) and CsrS (csrS) genes, complete cds. NID: g3599370.
 atgcttccaaacataaatgggtctagaaattttagacaaaattcgtaaaaaacaactact
 ccaattatcatcattactgcaaaaagcgagacatatgataaagtagctgggttgactat
 gggcgagatgactacattgtaaaaccctttgatatagaagaattgctcgcaagaataaga
 40 gcggtattgcgagacagccagataaagatgttttagatatcaatgggtattatcattgat
 aaagatgccttttaagttactgttaatggccatcaattagaattaactaaaacagaatac
 gatttattatatgttttagctgaaaatcgtaaccacgtcatgcagcgtgaacaaattctc
 gatcacgtatgggggtataatagtgaagtagaaacgaatgtcgttgatgtttacattcgt
 tatttacgtaataaactcaaacccttttaataaagaaaaatccatagaacagtagctggc
 45 gtagggtatgtgattcgatga

Sequence 2788

MLPNINGLEICRQIRQKTTTPIIIIITAKSETYDKVAGLDYGADDYIVKPFIDIEELLARIR
 AVLRRQPKDKVDLDINGIIIDKDAFKVTVNGHQLELTKEYDILLYVLAENRNHVMQREQIL
 50 DHVWGYNSEVETNVVDVYIRYLNRNKLKPFNKEKSIETVRGVGVIR*

Sequence 2789

Contig_0824_pos_1929_1075,

is similar to (with p-value 2.0e-42)

55 >gp:gp|U81166|LU81166_1 Lactococcus lactis subsp. cremoris
 MG1363 histidine kinase (llkinA) gene, complete cds. NID: g2
 182834.
 gtgagttatatcttttcttcgcaaattactaaaccgatagttacaatgtccaataaaatg
 aatcaaattagaagagatggttttcaaaaataaactgaattaactacaaattatgaagaa
 acagataatttaattgatacttttaataagaatgatgtatcaaatagaagaatcttttaatt

cagcaacgtcaatttgtcgaggatgcttcacacgaattaagaacgccactgcagattatt
 caaggtcatctaaatttaacccaacgttgggggaaaaaagatccagcagtttggagaa
 tctttgaatatttcaattgaagaagtgaatcgataacaaaacttgtcgaagaactactt
 ttacttaccaaagatagagtcaatcataatgttttggaaatgtgaaaatgtagacgtaaat
 5 agcgagattcaatcacgtgtgaagtcactgcaacacctacatccagattatactttttaa
 acacatcttgcactaagcctatccaattaaaaaataaccgcatcagtttgaacaactc
 ttactcatattttattgataatgcaatgaaatacgacactgaacataagcacattaaaatt
 gttactcaactaaaaaataaaatgattatgattgatattactgatcatggatgggtata
 ccaaaagctgacttagaatttatctttgatagattttatcgtgtagataaatcacgtgct
 10 cgtagtcaggaggcaatggattaggactatcaatagcagaaaaaattgtgcaacttaac
 ggtggtatgattcaagtagaaagtgaactacaaaagtacacgactttcaaaatcagtttt
 ccagtactaaactaa

Sequence 2790

15 VSYIFSSQITKPIVMTSNKMNQIRRDGFQNKLELTNYEETDNLIDTFNEMMYQIEESFN
 QQRQFVEDASHELRTPLQIIQGHNLNIQRWGKKDPAVLEESLNISIEEVNRITKLVEELL
 LLTKDRVNVHNVLECEVNDVNSEIQSRVKSLLHLPDYTFETHLATKPIQLKINRHQFEQL
 LLIFIDNAMKYDTEHKKHIKIVTQLKNKMIMIDITDHGMGIPKADLEFIFDRFYRVDKSRA
 20 RSQGGNGLGLSIAEKIVQLNGGMIQVESELQKYTTFKISFPVLN*

Sequence 2791

Contig_0824_pos_0_597,
 is similar to (with p-value 8.0e-29)
 >pir:pir|S25295|A32879 oxoglutarate dehydrogenase (lipoamide
 25) (EC 1.2.4.2) - Bacillus subtilis
 atgctagatttgtatgatgattatctacaaaatccatcatccgtacctgaagatttaca
 gtcttgttcagtacaattaaaacaggtgaagctcatatcgaagctaaacctaccactgat
 ggggggtggttcacaagcgggagatagcacaattaaacgtgttatgcgcttaatcgataat
 attcgtcaatacggacattttaaagcagatatttatccagtaaatcctccagagcgctcaa
 30 aatgttcctaaatttgaaatcgaagattttgatttagataaaagaaactttggaaaaata
 tcactcgtgaattgtctctgaacatttttaagacatttatgacaatgcctatgatgcaatt
 gttcgtatggaaagacgtttacaaaggaccgatagcttttgaatacactcacattaataat
 aataaagaacgtgtgtgtgttaaaagaagaattgaaacgccttataaagcaagtttaaac
 gataatcaaaaaaaagaacttttcaaaaaactcgacacgtagaagggttttgaaaaatat
 35 ttgcacaaaaattttgttgggctaaacgtttttcaattgaaggcgtcgaTC

Sequence 2792

MLDLYDDYLQNPSSVPEDLQVLFSTIKTGEAHEIAKPTTDGGGSQAGDSTIKRVMRLIDN
 40 IRQYHGLKADIYPVNPPEQNVKLEIEDFDLDKETLEKISSGIVSEHFKDIYDNAYDAI
 VMERRYKGPFAFEYTHINNNKERVWLKRRIETPYKASLNDNQKELFKKLAHVEGFKEY
 LHKNFVGAKRFSIEGVDX

Sequence 2793

Contig_0440_pos_5821_6999
 is similar to (with p-value 2.0e-24)
 >sp:sp|P23524|YHAD_ECOLI HYPOTHETICAL 42.1 KD PROTEIN IN RN
 PB-SOHA INTERGENIC REGION (ORF 3) (F408). >pir:pir|JQ0614|JQ
 0614 hypothetical 42K protein - Escherichia coli >gp:gp|D902
 12|ECORNPBW_3 E.coli rnpB gene and ORFs. NID: g216630. >gp:g
 50 p|U18997|ECOUW67_54 Escherichia coli K-12 chromosomal region
 from 67.4 to 76.0 minutes. NID: g606010. >gp:gp|AE000394|AE
 000394_2 Escherichia coli K-12 MG1655 section 284 of 400 of
 the complete genome. NID: g2367197.
 atgtttaaaataatttttggaaaagagaaaaataaggtggttaagacaatgaaagtttta
 55 gttagccatggatgaatttaattgaattatttctagttaccaagctaatagatatgttgaa
 gaagcggtagcaagtcaaattgaagatgcagatatcgttcaagttccactatttaacggt
 cgtcacgaattatttagattcagttctttctttggcaatcaggaaataaatatcgtgtgagt
 gcgcatgatgctgacatgaaagaaaccgaagcaatatatggacaaacggatagtggtatg
 actattatcgaaggtcacttatttttaaatggcaaaaaacctattcaacatcgatcaagt

tacggtttgggagaggttataaaagcagcattggacaatcatacagaacatcttgttatt
 tcttttaggtggaataggaagtttgcggtgcaggcatgttgcaagcattgggtgca
 acattttatgatgatgaagcacaattgtcgatatgaggaaagtgcatatttaataaaa
 tatattagacgtattgatttatcaggtgttcacccacaattacaaggtaaacattcaa
 5 ttaatgtcagattttctcaagtcgattgtatgggaaaaaaagtgaatcatgcaaacatac
 gaatcattagattttgtctcaaaatgaagcagccgagatagataatttaatttggtatttt
 agtgaattattttaagaatgaattgaaaatagcaatgggaccaatcgagcgcggtggtgct
 ggaggtggtatagcagctgtattaaatagctcttatcaagctgagattttaacaagccat
 gaattagtgaatcaaatcacacatttagaaaacttaattcaacaggcagatcttattatt
 10 ttccggagaaggtttgaagaagaagatcaaattctagagactacaacaatcgtatagca
 gaacttaccacaacaatcacagcaagccagctattgcaatttgcgtacaaatgataaattt
 gatttgtttgaatcattgaatgttacagcaatgtttaatacatttattgatatgcctgat
 tcatatacagatttttaagatgggtattcaaatcagacattacacagtacaagcactaaa
 ttattgaaaacgcaataaacttaccgctttcatcctaa

15

Sequence 2794

MFKIIIFGKEKNKVVKTMKVLVAMDEFNGIISYQANRYVEEAVASQIEDADIVQVPLFNG
 RHELLDSVFLWQSGNKYRVSAHDADMKETAIYGQTDGMTIIEGHLFLNGKKPIQHRSS
 YGLGEVIKAALDNHTEHLVISLGGIGSFDGGAGMLQALGATFYDDEAQIVDMRKAYLIK
 20 YIRRIDLSGVHPQLTKVNIQLMSDFSSRLYGKKSEIMQTYESLDLSQNEAAEIDNLIWYF
 SELFKNELKIAMGPIERGGAGGGIAAVLNSLYQAEILTSHELNVNQITHLENLIQQADLII
 FGEGLKEEDQILETTTIRIAELTQQYSKPAIAICATNDKFDLFESLNVNTAMFNTFIDMPD
 SYTDFKMGIIQIRHYTVQALKLLKTQINLPLSS*

25 Sequence 2795

Contig_0440_pos_4831_4487

is similar to (with p-value 1.0e-60)

>pir:pir|I67760|I67760 transposase (insertion sequence IS10
) - Escherichia coli >gp:gp|S67119|S67119_2 BST=somatotropin
 30 ...BST/beta-Gal fusion protein [Escherichia coli, LBB84, pla
 smid pXT107, IS10L/R-1, PlasmidTransposonInsertionMutant, 3
 genes, 1679 nt]. NID: g455674.
 atgcagattgaagaaaccttccgagacttgaaaagtcctgcctacggactaggcctacgc
 catagccgaacgagcagctcagagcgttttgatatcatgctgctaategcacctgatgctt
 35 caactaacatgttgcttgcgggcgttcatgctcagaaacaagggttgggacaagcacttc
 caggctaacacagtcagaaatcgaaacgtactctcaacagttcgcttaggcattggaagtt
 ttgcggcatttctggctacacaataacaaggaagacttactcgtggctgcaaccctacta
 gctcaaatattattcacacatggttacgctttggggaaattatga

40 Sequence 2796

MQIEETFRDLKSPAYGLRLHSRTSSSERFDIMLLIALMLQLTCWLAGVHAQKQGWDKHF
 QANTVRNRNVLSTVRLGMEVLRHSGYTTITREDLLVAATLLAQNLFTHGYALGKL*

Sequence 2797

45 Contig_0441_pos_2559_3467

is similar to (with p-value 2.0e-39)

>sp:sp|054983|CRYM_MOUSE MU-CRYSTALLIN HOMOLOG. >gp:gp|AF03
 9391|AF039391_1 Mus musculus mu-crystallin (Crym) mRNA, comp
 lete cds. NID: g2745895.
 atgaatgaagtcattttagaagtagaaaaagctttgcaagctttttcagagaataagacg
 ataacgccattaaagatatgttttgcatttaattagcaaaatcgattatttagtgatgcc
 50 ccatattcagatgaattaaatcgttggaacttaaaacagtcctcatttgcacctgaaaat
 tcaaaaaaagggaagcgactattactggatcagttattttaagtactatgaaacagga
 gaaacattgtctatattagatgggtgttttcttactaaagtaagaactggtgcaatttca
 55 ggtgtagctactaaatatctagcaaaagaaaacgctaaacacttagtgtaataggggca
 ggtgtacaaagctgaaggtttaattgaagcgatacttgctggttagagatattgaaaaaatt
 cacatcgcaagtagaacgttcgaaaaagcagaaaaatttgctcaaaatatacgaaatcga
 tttaatattaaagtgaagtgtatttagatcggcagatgaagcgatagacagtgagatatt
 gtagttacagcaacaatgcaaatcagcccgctttatactcattctttacatccaggcgtg

catttaaattgcagtcggatcctttaaacagatatgcaagaaatccttcagaaacaatg
 cttgttgctaataaaattgttgaatctatggaagcagctttagaagaaacaggtgat
 ttaaaaattcctcaagcagaaggaatattaactaaaaatagctacatagtgaattaggc
 gacattatttctggtgaaaaagttggccgcgaactgaagaagaagtgacagctctttaa
 5 tcggtcggctctagcaattgtagatatcattgtggcacaatattttataaaaaattaata
 caatcttag

Sequence 2798

MNEV:LEVEKALQAFSENKTITPLRYVLPFNEQNRYLVMPALSDELNIVGLKTVSFAPEN
 10 SKKGKATITGSVILSDYETGETLSILDGGFLTKVRTGAISGVATKYLAKENAKTLSVIGA
 GVQAEGLIEAILAVRDIEKIHASRTFEKAEKFAQNIRNRFNIKVSVFRSADEAIDSADI
 VVTATNANQPVYTHSLHPGVHLNAVGSFKPDMQEIIPSETMLVANKIVVESMEAALLETGD
 LKIPQAEGLITKNMLHSELGDIISGEKVGRETEEEVTVFKSVGLAIVDIIVAQYFYKKLI
 QS*

Sequence 2799

Contig_0441_pos_4023_4979
 unknown

atgaaaaaataataatcactggagcacttggtcaaattggtactgaattagttattaag
 20 tgtagagaaagatatggaactgaaatgtattagctactgatataagaaagccagaacca
 cactctccagttaaaaacggtccttttgaaatcttagatgttacagatagaaatcgttta
 tttgaaacagtgagatatcattcaatgctgatacacttatgcataatggctgcactgcttct
 gctacagcggaaaaagaaaccactagttgcctgggatttgaatatggggggacttataaat
 acactcgaagcagcaagaaggtatcaattaaaaatattttacgccaaagttcaataggagct
 25 tttggtatttcgacacctaagtaaatcacctcaattaacaatacaacaaccaacgact
 atgtatgggtataaacaagaactgagggaattattatgtcagttattattatgtcaaattt
 ggagtagatacacgtagcgttagatttctggtttaatttcacatgttaaagaacctggc
 ggtggtacataacagattatgcagttgatatttttaaggcagttagaaaaggacattat
 acaagttatattaatcggtacacttatatggatatgatgtacatggaagatgctaitgat
 30 gcgattatcaaattaatggaagaagacagtgtaaattaaaaacaagaaatggctataac
 ttaagtgcgatgagttatgaacctgaaatgctaaaaacaagctatacaagtatactatcca
 gatttcaccttagattatgacattgatttagagcgtcaagatattgcattaagttggcca
 gatagtatagatacagagctgtgctcaggaagagtggggatttgatcctaaatatgattta
 ccaactatgactaaagtgtatttagaggcaattgagaaaaaacaagaatgctga

Sequence 2800

MKKIIITGALGQIGTELVIKCRERYGTENVLATDIRKPEPHSPVKNPFEILDVTDNRNL
 FETVRYFNADTLMHMAALLSATAEKKPLVAWDLNMGGGLINTLEAARRYQLKYTFPSIGA
 FGISTPKVNTPLTIQQPTMYGINKVTGELLQCYYYVKGVDTRSVRFPGLISHVKEPG
 40 GGITDYAVDMYFKAVRKGHYTSYINRYTYMDMMYMEDIAIDAIKLMEEDSVKLKTRNGYN
 LSAMSEPEMLKQAIQVYYPDFTLDDYDIDLERQDIALSWPDSIDTSCAQEEWGDFPKYDL
 PTMTKVMLEAIEKKQKEC*

Sequence 2801

Contig_0442_pos_1526_2359

is similar to (with p-value 1.0e-60)

>pir:pir|I67760|I67760 transposase (insertion sequence IS10
) - Escherichia coli >gp:gp|S67119|S67119_2 BST=somatotropin
 ...BST/beta-Gal fusion protein [Escherichia coli, LBB84, pla
 50 smid pXT107, IS10L/R-1, PlasmidTransposonInsertionMutant, 3
 genes, 1679 nt]. NID: g455674.

atgaagagtattttttaaagtaataatgtgaaaatattcacaaaaaaactaggagga
 tttgctatgagtaagaaaatattcgatacttttaaatataatgtggtgccaattaaaa
 aatagagtattaatggcaccatgactatccaagctgggtattttgatggaagtgttaca
 55 tcagaaatgattgattattatcaatttagagctggtgatgcttcagcaatcattgttgaa
 agttgttttgggtgaaaatcacggacgaggatttccgggagctataggtattgataatgat
 gacaaaaatcctggactcaaacgtttagcagaagcgattcaagctaagggatcaaaagcg
 attttgcaactttatcatgccggaagaatggcaatcctaaatttaatgaaggagagcag
 ccgatattctgcgagccccattgcagcattaagacctgatgctgtaccacctagagaaatg

acacatgctcaaatcaatcagatgattgatgactttggagaggctacacgtcgcgtata
 gaagcggggtttgatggtgtcgaaattcatggcgccaacacatacttattacaacaattt
 ttctctccacattctaatacggagacaagattcatggggaggcagtcgtgaaaaacgtaca
 cgatttccaatcgaagttttgacaaaggttcaacacgtcgttgctgaaaaagaggcttct
 5 cattttattatagatatcgattctcacctgaagaattgaagaaccaggcatatcgtttt
 gaagataccatgtttttactaaatacattagcagaatgcttgactacaaaatag

Sequence 2802

MKSILFLSNNVKIFTKKLGGFAMSKEIFDTFKFKCGAELKNRVLMAPMTIQAGYFDGSVT
 10 SEMIDYYQFRAGDASAIIVESCFVENHGRGFPGAIGIDNDDKIPGLKRLAEAIQAKGSKA
 ILQLYHAGRMANPKFNEGEQPIASPIAALRPDAVPPREMTHAQINQMIDDFGEATRRAI
 EAGFDGVEIHGANTYLLQQFFSPHSNRRQDSWGGGREKRTFPIEVLTQVQHVVAEKEAS
 HFIIGYRFSPEEIEEPGIRFEDTMFLNLTLAECLTK*

15 Sequence 2803

Contig_0442_pos_13117_13461

is similar to (with p-value 5.0e-53)

>sp:sp|P54524|YQIG_BACSU PROBABLE NADH-DEPENDENT FLAVIN OXI
 DOREDUCTASE YQIG (EC 1.-.-.-). >gp:gp|D84432|BACJH642_230 Ba
 20 cillus subtilis DNA, 283 Kb region containing skin element.
 NID: g2627063. >gp:gp|Z99116|BSUB0013_132 Bacillus subtilis
 complete genome (section 13 of 21): from 2395261 to 2613730.
 NID: g2634723.

atgcagattgaagaaaccttccgagacttgaaaagtcctgcctacggactaggcctacgc
 25 catagccgaacgagcagctcagagcggtttgatcatgctgctaatacgccctgatgctt
 caactaacatggttgcttgcggcggttcagctcagaaacaaggttgggacaagcacttc
 caggctaacacagtcagaaatcgaaacgtactctcaacagttcgcttaggcattggaagt
 ttgcggcattctggctacacaataacaagggaagacttactcgtggctgcaaccctacta
 gctcaaaatttattcacacatggttacgctttggggaaattatga

30

Sequence 2804

MQIEETFRDLKSPAYGLGLRHSRTSSSERFDIMLLIALMLQLTCWLAGVHAQKQGWDKHF
 QANTVRNRNVLSVRLGMEVLRHSGYTITREDLLVAATLLAQNLFTHGVALGKL*

35 Sequence 2805

Contig_0445_pos_0_322

is similar to (with p-value 8.0e-17)

>gp:gp|M89774|ECOLYSP_2 Escherichia coli lysine specific pe
 rmease (lysP) gene, complete cds. NID: g466776. >gp:gp|AE000
 40 305|AE00305_1 Escherichia coli K-12 MG1655 section 195 of 4
 00 of the complete genome. NID: g1788479.

gtggcaactggaagtgtcatttctcaagctggcccaggaggagctatattagcttatata
 ctaattggtattatgctttattttttaatgtcatcaataggagaattagcaactttctat
 ccggtttctggttcttttagttcatactctaccagatttgttgattcgctcacttggtttt
 45 acaatgggttggttgattggggtatgtggtcacttgtaacaagtgtagatatcattggt
 gcttccaatgtattacaatattgggatgtatttaaagttttaaatccacttacatggagc
 ttaattttcttaactctgttgt

Sequence 2806

50 VATGSVISQAGPGGAILAYILIGIMLYFLMSSIGELATFYPVSGSFSSYSTRFVDSSLG
 TMGWLYWGMWSLVTSVDIIVASNVLQYWDVFKVLNPLTWSLIFLTLX

Sequence 2807

Contig_0447_pos_14098_15573

55 is similar to (with p-value 6.0e-40)

>gp:gp|X93084|MBFMDSUBS_1 M.barkeri fmdE, fmdF, fmdA, fmdC,
 fmdD, fmdB, orf4, orf3, orf2, and orf1 genes. NID: g1124956

atgaaagattgattactcttattgtcatgatatcttttgttttagcgagctgtggggc

acatcaagtacagacaaagacaccctcaatggtgaaatacctttgaaaactaaatcaatt
 gcaccttacgaaactgatatcccaggtaaaaacaggtgccttggaatcgctttttaaaatg
 tcgaagaatggtaaagtaaaacctttatttagtcaaaaattatcatcaagtatctgacaat
 caactagaactcactttaaaagataatattaaattcaaaacggtcatcatttaacaggc
 5 gaagctgtaaaacgtagtctcgaagaaggaatgaaaaaagtgatttggtaaaaggatca
 ctctctattaaatcaatcaatgctcatggacaaaaagtcacaatcactactaaagaacct
 tatccagaattaatgtctgaactcgcaagcccatttgctgctatttacgacacaaaagct
 aaaaacaaagtaactgatcaacctgttggtacgggtccttataaaattgatcagtataaa
 cgttcgcacaaaaatcgtactaaaacaattcaaaagactactggcaaggtacgccaaaatta
 10 aaaagaattaatgtcacttatcatgaagatggaataactcgtgttgatcacttattatca
 ggcaaatcagatttgactactgatgttccaattgaacgcgttgatgatgtaaaaaaatct
 aacaaagcaaacattcaaaagtacatcaggctttagaacgcatttaattgttatacaatcat
 gatagtaaaaaagttataaaaaagtaagagaagcactagatatgattattaatcgaaaa
 gacattgctaaaaatgtttctaaaaattatgctgagccagcatcagggtccttttaacct
 15 cgattaaaaatcattagaaaaagaggaaattcaatcacaagacatcaagcgtgcaaaagaa
 ctttttagctcaagaaggttattccaaatcgcatcctcttaaatataacatggtcacatac
 gatggcagaccagaattgcctaaaattggacaggtgatacaatctgaagctaaaaaggca
 aatggttgatatacaattacgcaatgtagatgatcgaaggatatctcaaaaacaaacag
 agttgggatgtttcaatgtatagttatttaagtgtgccacgtggtgatacaggttatttc
 20 ttttaacactgcatacttacctgatggagcattaaataaaggtaattatagtaataactaaa
 gtcactcagtttaattaaagaattaaataactactttcggtgacaagcaacgtgggtcaagt
 actaatgaaataactaaatgaatctaaaaaagatattcctaacagctatatcacatacaac
 tctcaaatagatgggtgtgaataataaagtaagacattttaattgttacaccagaatctatc
 25 tatttaattgattataaattaaagtaaaaaagaataa

Sequence 2808

MKKLITLIVMISFVLASCGGTSSTDKDTLNV EIP LKTKSIAPYETDIPVKTGALES LFKM
 SKNGKVKPLLVKNYHQVSDNQLELTLKDNIKFNQNGHHLTGEAVKRSLEEGMKKSDLLKGS
 LPIKSINAHGQKVTTITKEPYELMSELASPFAAIYDTKAKNKVTDQPVGTGPYKIDQYK
 30 RSQKIYLVKQFKDYWQGT PKLKRINVTYHEDGNTRVDHLLSGKSDLTTDVPIERVDDVKK
 NKANIQTSGFRTHLMLYNHDSKKVNKKVREALDMI INRKDIAKNVSKNYAEPASGPFNH
 RLKSLEKEEIQSQDIKRAKELLAQEGYSKSHPLKLNMTYDGRPELPKIGQVIQSEAKKA
 NVDIQLRNVDDIEGYLKNKQSWDVSMSYLSVPRGDTGYFFNTAYLPD GALNKGNYSTK
 VTQLIKELNTTFGDKQRGQVTNEILNESKKDIPNSYITYNSQIDGVNNKVRHFNVTPEST
 35 YLIDYKLSKKE*

Sequence 2809

Contig_0447_pos_20772_21911

>gp:gp|282038|CT282038_4 C.thermosaccharolyticum etfB, etfA
 40 , hbd, thlA and actA genes. NID: g1667352. >gp:gp|292974|TTB
 CSOPRN_6 T.thermosaccharolyticum BCS operon DNA. NID: g19033
 26.
 gtgttttggtggtgtattttaaggatatacctgcctatgaactaggtgcaacagttattcgt
 caaattttagaacatagtc aaatagatcctaataaagaaatgaagttattctaggaaac
 45 gtattacaggcaggtcaaggacaaaatcctgctcgatttgctgcgattcatgggtggtg
 ccagaagcgggtaccttcttttactgtaaataaagtttgcggttctggattaaaagcgatt
 caacttgctatcaatctattgttagcgggagataatgagattggtatcgctggaggcatg
 gaaagtatgtctcaatctccaatgcttcttaaaaatagtcgttttcggttttaaaatggga
 aatcaaaactttagaagatagtatgatagctgatgggttaactgataagtttaattgattac
 50 catatgggtatcacagccgaaaaatctagttgaacagtatcagattagtcgtaaagaacaa
 gatcaatttgcatctcgattctcaacaaaaagcatcacgtgcacacaagctgggtgattt
 gatgctgaaattgtacctgtagaggtaccacaacgtaaaaggcgaccccttaattatttct
 caagatgaaggcattagacctcaaacgacaattgataagttagcacaactccgtccagca
 tttaaaaaagatggatcagtaactgctggtaatgcatccggtatcaatgacggtgctgct
 55 gctatgctcggttatgacggaggacaaagcgaaagcattgggcttacaacctatagctgta
 ttatagatagttttggtgcgagtggtgtggcgcttcaattatgggtattggtccagttgaa
 gcgatacataaaagctttaaaacggttctaataaagtataaatgatgttgatatttttgaa
 ttaaacgaagcttttgacgcgaatcaattgctgtaaacggtgagttgcaattaccgcaa
 gataaagtcaatgttaattggtggtgcgattgcactaggacatccgataggggttcgggt

gcgcgtacttttagtttcattattacatcaattaagtgatgctaagccaacaggtgtggca
tctttatgtatcggtggcggtcaaggtatcgctacggttgatctaaatatgaagtttaa

Sequence 2810

5 VFGGVFKDIPAYELGATVIRQILEHSQIDPNEINEVILGNVLQAGQGQNPARIAAIHGGV
PEAVPSFTVNKVCGLKAIQLAYQSIVAGDNEIVIAGGMESMSQSPMLLKNSRFGFKMG
NQTLEDSMIADGLTDKFNDYHMGITAENLVEQYQISRKEQDQFAFDSQOKASRAQQAGVF
DAEIVPVEVPQRKGDPLIISQDEGIRPQTTIDKLAQLRPAFKKDGSVTASNAGSINDGAA
10 AMLVMTEDKAKALGLQPIAVLDSFGASGVAPSIMGIGPVEAIHKALKRSNKVINDVDIFE
LNEAFAAQSI AVNRELQLPQDKVNVNGGAI ALGHPIGASGARTLVSLHLQLSDAKPTGVA
SLCIGGGQGIATVVSKYEV*

Sequence 2811

Contig_0447_pos_24387_22894

15 >pir:pir|S57636|S57636 5-methyltetrahydropteroyltriglutamat
e--homocysteine S-methyltransferase (EC 2.1.1.14) - Madagasc
ar periwinkle >gp:gp|X83499|CRMETS_1 C.roseus MetE mRNA for
methionine synthase. NID: g886470.
gtggcggttttaggattagatttgggtacacgataacggctataacttaaaacaaattgaa
20 gatggtaatttcgatcaaaagtaaagcactttatgcaggaatcattgatggcagaaatgta
tggcggtgatattgaagcaaaaaacaattaatcgaaacattacaacaacacacacaa
cagtttagtcattcaaccttcacatcacttctacatgttccagtatcactagatgatgaa
acacttgatgaatctattgcagaaggattaagtttggctactgaaaagttagacgaatta
gatgcattgcgtcgcttattcaatgataatgacctgtcaaaatgaacattacaagca
25 cgttatgaacgcttccaaagccaatcatttaaaaatcttgaatcagattttgaaagtga
ccgacacatcgtaaatcaccatttgcataacgcaagcaattacaaaaccaacgtctaaat
ttaccagatttacctacgactacgatttggttcattccctcaaacctcgtagaagtcgtaaa
tttagagcagattggaaaaataatcgccattacagatgctgagatcaagaatttttaca
aatgaaatcgctcggttgattaaaaatccaagaagatattggtttagacgtagctgtacac
30 ggtgagtttgagcgtaacgacatggttgaatttttggggaaaaacttcaaggtttccta
gtaacaaaattcggttggttacagtccttacggttctcggtcagttaaaccaccagttata
tatggagatgtgaagtggactgccccacttactgtaaaagaacagtatagcgcaaaagc
ttaactgacaaaccagtaaaaggatgcttacaggtccagttactattttaaactggtca
tttgaaacgctagacgtacctaagaagggtgttcaagatcaaattgccttagccattgat
35 gaagaagtgttagctttagaagaagcaggtatcaaagtcattcaagttgacgaaccaacg
ttgcgtgaaggcttacctctacgctctgaatatcacgaacaatattagaagatgctgta
cattcatttaaattagcgacatcatctgtgcatgatgaacacaaattcatacacatatg
tgctattctcaatttggaacaaatcattcatgctatccatgatttagatgcagatgttatt
tctattgaaacttctcgtagtcacggagacttaattcaagatttcgaagatattaattat
40 gatttaggaattggattaggtgtttacgatattcacagtcacagttattccaaccgaagaa
gagattactactgcaattaatcggttcattacaacaaattgatcggttctctattctgggtt
aacccagactgtggtctaaaaacacgtaaaagaaatgaagtgaagatgcattaaacagtg
ttagtcaatgcagttaaaaagaaacgtcaagaatctgaatcaacaacagcataa

Sequence 2812

45 VGGLGLDLVHDNGYNLQIEDGNFDQSKALYAGIIDGRNVWAADIEAKKQLIETLQOHTQ
QLVIQPSSSLHLHPVSLDDETLDESIAEGLSFATEKLDELALRRLFNDNDLSKYEHYKA
RYERFQSQSFKNLEYDFESVPTHRKSPFAKRKQLQNQLNLPDLPTTTIGSFPQTREVRK
FRADWKNNRITDAEYQEFQNEIARWIKIQEDIGLDVLVHGEFERNDMVEFFGEKLQGF
50 VTKFGWVQSYGSRVKKPPVIYGDVKTAPLTVKETVYAQSLTDKPVKGMMLTGPVTILNWS
FERVDVPRKVVQDQIALAIDEEVLALAEAGIKVIQVDEPTLREGLPLRSEYHEQYLEDV
HSFKLATSSVHDETTQIHTHMCYSQFGQIIHAIHDLADVISIETSRSHGDLIQDFEDINY
DLGIGLGVDYDIHSPRIPTEEETITAINRSLQQIDRSLFWVNPDCGLKTRKENEVKDALTV
LVNAVKKKRQSESTTA*

Sequence 2813

Contig_0450_pos_7632_8510

No hits found

atgaaagcgccagttctggtatcaggtactgacggtgtgggtacaaagttaaaattagca

attgactatggaaagcatgacacaattgggtattgatgctgtcgcaatgtgtgtaaataatgat
 attttaacaacaggtgctgaacctttatacttttttagactatattgccacgaataaagta
 gtgccaagtactatagagcaaatcggttaaaggtataagtgacggttgcgaacaaaccaat
 acggcacttataggcgggtgaaactgctgaaatgggagaaatgtatcatgaaggtgaatat
 5 gatattgctggttttgcagtaggagcggtagagaaagaggactatattgatggttcaa
 gttgaagaaggacaagcaattattggttttagcttcaagtggattcattcaaattggctat
 agtctagttagaaaaatgataaaagaatcaggagttcaattacatgatcaatttaattggt
 caaaccttttttagaaaccttccttgcaccaacaaaattgtatgtaaagcctattcttgaa
 10 ttaaagaaacataattgatatacaagcgatgagccatattactggtggagggtttctatgaa
 aatattccgcgtgcccttcctaaagggttatcagcaaaaatagatacacaatcattccca
 acggttgaagtcctttaattggcttcaaaaacagggcaacatttcaacgaatgaaatgtat
 aacataatttaattgggtattggatatacaattattgttgacaaaaaagatgttcaaaaa
 acattaacaacggttacgtgcaatggatacaactgcatatgaaattggtgagattataaaa
 gatgatgatacacctattcatttattggaggtagaataag

15

Sequence 2814

MKAPVLVSGTDGVTGLKLALIDYKHDITIGIDAVAMCVNDILTGAEPFLYFLDYIATNKV
 VPSTIEQIVKGISDGCETNTALIGGETAEMGEMYHEGEYDIAGFAVGAVERKEDYIDGSN
 VEEQAIIGLASSGIHNSGYSLVRKMIKESGVQLHDQFNGQTFLETFLAPTKLYVKPILE
 20 LKKHIDIKAMSHITGGGFYENIPRALPKGLSAKIDTQSFPTELVFNWLQKQGNISTNEMY
 NIFNMGIGYTIIVDKKDVQTTTLTLRAMDATTAYEIGEIIKDDDTPIHLLEVE*

Sequence 2815

Contig_0450_pos_13563_14528

25 is similar to (with p-value 9.0e-27)

>gp:gp|Z99108|BSUB0005_100 Bacillus subtilis complete genom
 e (section 5 of 21): from 802821 to 1011250. NID: g2633055.
 >gp:gp|D78508|D78508_5 Bacillus subtilis DNA for YfiO, YfiP,
 YfiN, YfiM, YfiL, YfiK, YfiJ, YfiI, YfiH, complete cds. NID
 30 : g1817531.

gtgggttaaaggcaacaatcttggcaaaaatgctgctctcaaaacttataatgagaatcaa
 gcgcaacatctgcttaaacagaatcaattacaaggatactttgtcttcgaccgtggtatg
 accgatactttttataaagatggttagccttcctataactatttatacatatgatgaacaa
 tcaagtaacagtgctggtggttaatacaattaacacgctcagtttatgaccgttttaagtta
 35 tcaatggcggtgtgctaagctttaatacaattagctaaagatccttcaaataagacgta
 gcaatgacattgattgatattgtattaccggtttaaatcggttcaggttcatttaatttt
 gaaccatacatatttatgacaccagtagttattatgtagtcactggattttctttgtct
 atctttatattgtgtttatcactttatacagtactaaaaatgaatcaagaaactgcactc
 aaagaacgcttgcaaatgtttcatttttccttggaaaagctcacgatagttcgaggtatc
 40 attgcatgggtttttactcactcatatgggcatttattggcctttatctggattactcatgct
 ctaaatgccccatttgaaaaatacaattggccaacggttagctttacaactcacttattac
 gttacgttactcgctcctatgcttattacttatagacttaattactcggttcattgataaac
 tttctactcaaattattacttagctttgttatcggtatttttctgggataaattccct
 actatcttctttaaacacatgcttaattgatgtaatacattacacaaccatttagtttggtt
 45 actaatcaaatgttagaaataacactcaataactatatttttagacacacatccagcattt
 tatttaagttttattacactattgatactattcatcattgttttagtatggaggtagcgc
 cgatga

Sequence 2816

50 VGKGNLGLKNAALKTYNENQAQHLKQNLQGYFVFDGRMTDTFYKDGSLPITIITYYDEQ
 SSNSVVVNQLTRSVYDRLMLSMGGVLSFNQLAKDPSNEDVAMTLIDMLFTGLNRSFSNF
 EPIHIYDTSSYYVVTGFLLSIFILCLSLYTVLKMNQETALKERLQMFHFSFEKLTIVRGI
 IAWFYSLIWAFIGFIWITHALNAPFEKYNWPTVALQLTYTYVTLVLCLLLIDLITFSWIN
 FLLKLLISFVIVIFSGIIIPITIFFKMLNDVITQPFSLVTNQMLEITLNNYILDHPAF
 55 YLSFITLLILFIIVLVWRYRR*

Sequence 2817

Contig_0450_pos_15460_16149

is similar to (with p-value 1.0e-18)

>sp:sp|P22082|SNF2_YEAST TRANSCRIPTION REGULATORY PROTEIN SNF2 (SWI/SNF COMPLEX COMPONENT SNF2) (REGULATORY PROTEIN SWI2) (REGULATORY PROTEIN GAM1) (TRANSCRIPTION FACTOR TYE3). >pir:pir|S15047|S15047 SNF2 protein - yeast (*Saccharomyces cerevisiae*) >gp:gp|X57837|SCSNF2_1 Yeast GAM1/SNF2 gene for a nuclear protein required for transcription of STAI gene. NID: g4499. >gp:gp|X89633|SCVCOSMGN_20 *S.cerevisiae* DNA for VPH1, MOD5, CAP20, ORF1 and SNF2 genes. NID: g1279694. >gp:gp|Z75198|SCYOR290C_1 *S.cerevisiae* chromosome XV reading frame ORF YOR290c. NID: g1420643. >gp:gp|D90459|YSCRIC1_1 Yeast RIC1 gene (regulatory gene for phospholipid synthesis), complete cds. NID: g806531. >gp:gp|M61703|YSCSNF2A_1 *Saccharomyces cerevisiae* SNF2 protein gene, complete cds. NID: g172631.

atgatagatatattcaaaatgtttccaaaagctataaaaagaagcatattttcgattcccta
 gatatgcaatttcaaaatcataaaattactattttattaggtgaaaatggtgctggaaaa
 tctacattattgctgttaattgcaggtattgagaatgcagacgaaggacgtattcaatac
 ttcaatcaatattttgtcaagacgtcgaatacgtcatattgttaggtatgtccctcaagac
 atcgactatttcgagcatatgactgtcatggagaacattgagttttcaagtcactttgt
 gaaaatcctatttcagatgaaacacttctattttatcacaaattaaattttactgat
 acaaaagtgaagtatctaacctttctgggggaaataaacgtaaaagtcaatattatgata
 ggtctacttagtcggcctaaaatacttattctagatgagccaacagaaggcattgattta
 gaatcaagatatgatattcacaacttattacaacaaatgaccgatcaatgtttaatcatc
 atgacgcacacatcatttagacgaagttgaagcactagcagatgatattaaagttataggt
 caaaatcctttttatcatgatattttagaaaataaaggttggtcttttaaaaaaatatgca
 aatgccttagctgataatacgaatcttaa

Sequence 2818

MIDIQNVSKSYKKKHIFDSLDMQFQNHKITILLGENGAGKSTLLRLIAGIENADEGRIQY
 FNQYLSRRIRHIVGYVPQDIALFEHMTVMENIEFFKSLCENPISDETLHSYLSQLNFTD
 TKVKVSNLSGGNKRKVNIMIGLLSRPKILILDEPTEGIDLESRYDIHNLQQMTDQCLII
 MTHHLDEVEALADDIKVIGQNPFYHDILENKGWSFKKYANALADNTKS*

Sequence 2819

Contig_0450_pos_20344_21786
 is similar to (with p-value 5.0e-85)
 >sp:sp|P43848|PUR5_HAEIN PHOSPHORIBOSYLFORMYLGLYCINAMIDINE
 CYCLO-LIGASE (EC 6.3.3.1) (AIRS) (PHOSPHORIBOSYL-AMINOIMIDAZ
 OLE SYNTHETASE) (AIR SYNTHASE). >pir:pir|G64122|G64122 5'-ph
 osphoribosyl-5-aminoimidazole synthetase (purM) homolog - *Ha
 emophilus influenzae* (strain Rd KW20) >gp:gp|U32822|U32822_2
Haemophilus influenzae Rd section 137 of 163 of the complet
 e genome. NID: g1574265.

atgaaaaagagcagccaatttttctgtacctggagctggtaaaacggctatgatgtatggc
 acatttgctttttgtctagtgaataaaagcggaaagttgataaattaatagttatttct
 ccaattaatgcatttgaagcttggcggttcagaatttattgaagtttttcaagataaaaga
 gaattacactttatgaacctaaagagataaaaaatataatgatttaggtaaaagtacgaaca
 gattggggaagtgcaaatgtcattgttttgaattttgaagcaatacaaaagtatgtaggg
 gttttaaatgaacttattaatgataagacaatgatagtttatgatgaggttcataggata
 aaaggtattaatagtagcagagcaagttatgcattaacttttaggtcctaaaagttattac
 agatacgttttaactggtaccccaattccaaatagttatcaagacatatttaacttctta
 aatattttatataaagatgagtatgatacttattttggttggaatgttgctgatttaca
 aatcctgatcctaatagaattaatgacaagttgaaccctttttctggcgtaacaataag
 aacgatttgggaagtcctcaagcagaaaatgatattatattatgtgtaagcctagta
 attcaaatgaattagcaaaagcgatatacgaatgaatctgggatactagcgatttat
 ataaggttactacaagcttcaactaaccagaattattgcaaaagaatattaattatagc
 gaactaggaatgttgatgaattaaatttcgatttgataaagcattaaataaagaa
 gaagaaaatcaagtaaaacaacaaatttataattcttttgatttaaaaaatgtaacttct
 ccaaaattcgaaaaaggtattgaattgattgaaaacttagttagccaaggaaaaaaggt
 ctagtatggggaattgtttgtagggtacaatgaataaaatcaataagaggttactagaaagt

gatattaattcaatattgatttatggagaaacacctaagaagatagggtagatatgac
 aataatttttaggaatggaatgcacaagttctaataatcctaatacattaggcgag
 tccatatctttacatcagacagtagatgcaatatattttgaatataactttaattta
 acgtttatgttgcaatcacgtgtagaatacatcgtttagggtaaataataatcaatat
 5 acaagggtattattttgatgtctgaaggggtagagcccataaagggttttatcgataaa
 gcagtttataacagactgaagggaaaagaagatgtaatgttaaatgctattgatggaat
 actttaagccaatgattgaagatgattacttagaagatgttaagaaaattattattgaa
 tga

10 Sequence 2820

MKRAANFSVPGAGKTAMMYGTFAFLSSEIKRKVDKLIVISPINAFEAWRSEFIEVFQDKR
 ELHFMMNRDKKYNDLGKVRTDWGSANVIVLNFEAIQKYVGVNLINDKTMIVYDEVHRI
 KGINSSRASYALTLPKSYRYVLTGTPIPNQYQDIFNFLNLYKDEYDTYFGWNVADLQ
 NPDPNEINDKLNPFWRNKNNDLEVPQAENDIILCVKPSNIQIELAKAIYENESGILAIY
 15 IRLLOASTNPPELLQKNINYSGLMLNDELNFDLKDALKNEENQVKQQIYNSFDLKNVTS
 PKFEKGIELIENLVSQGGKVLVWGLFVGTMMNKINKRLLESINSILYGETPKEDRVDMI
 NNFRNGNAQVLIISNPNTLGESISLHQTVHDAIYFEYNFNLTFMLQSRDRIHRLGLVNNQY
 TRYYYLMSEGDRAHKGFIDKAVYNRLKEKEDVMLNAIDGNTLKPMEIDYLEDVKKIIIE
 *

20

Sequence 2821

Contig_0451_pos_806_1621

is similar to (with p-value 1.0e-20)

25 >gp:gp|U67964|EVU67964_2 Ectromelia virus H14-B and H14-E g
 enes, complete cds. NID: g2145123.

atgcattatataaaatttattgagtcaaaagataatacaaaaactttatatgaaagtgaat
 gatattcaagatgcaaaagcgaatatcattatagctcatggtgtggcagaacatttagat
 cgttatgatgagataacagcatatttaaatgaagcgggttttagtggtatttagatatgat
 caaagaggggcatggtcgttctgaaggcaagcgtgccttttatagcaatagtaaatgaaatt
 30 gtcgaagatttagatgagataataaattatgtgaagtcaaactttgaaggtaaaagtttac
 ttaattcgggtcatagtatgggtggttatacagtcactttatatggaacgaaacatccaaat
 acagtgaaatgggtattataaacttctggagcattaacacggttataataataaactatttggc
 aatcctgatagaaacatatcacctgatacttatatagaaaacaatttaagtgaaggggta
 tgttctgatttagaggaatggaaaaatataaacttgatgatttgaatgcgaaacaaatc
 35 tctatggggctcgtctttcaataatggatgggtggttaggtatttgaaagacaatgctcaa
 caatttacagataatattttgatattgcattggcaaggaagatgggctagtaagcta:gtg
 gattctt:acagctttatcaagaaataggatcagcacataaatcattacacatctatgat
 cgtttgagcatgaaatatttaaatgaaagttcttataatagaactatttttaacgaagtt
 attgaatggcttgaaacggaattaacttataactaa

40

Sequence 2822

MHYIKFIESKDNTKLYMKVNDIQDAKANIIIAHGVAEHLDRYDEITAYLNEAGFSVIRYD
 QRGHGRSEGKRAFYSNSNEIVEDLDAIINYVKSNEFGKVYLIGHSMGGYTVTLYGTKHPN
 TVNGIITSGALTRYNNKLFGNPDNRNISPDTYIENNLSEGVCSOLEVMEKYKLDLNAKQI
 45 SMGLVFSIMDGVRYLKDNAQQFTDNILILHGKEDGLSVYVDSLQLYQEIGSAHKSLHIYD
 RLEHEIFNESSYNRTIFNEVIEWLETELTYN*

Sequence 2823

Contig_0455_pos_4830_4159

50 is similar to (with p-value 4.0e-17)

>sp:sp|Q10092|YAOD_SCHPO HYPOTHETICAL 24.2 KD PROTEIN C11D3
 .13 IN CHROMOSOME I. >gp:gp|Z68166|SPAC11D3_13 S.pombe chrom
 osome I cosmid c11D3. NID: g1107889.

gtgtatatattcatgagtaaaaaagttttatttgttttaacaagtacaagtcaatttaca
 55 gacggtacagaaactggattatgggttagaagaagctggagcaccatataatatattgact
 gaaagaaggtatcaatggttgatgttatttctattaaagtggaagtaaatcttgatccct
 aattctgtrttctaataatcactgaatcagtatgctaaattcgtgtcacacttaaagcat
 acacctagtatcgaaaatgtaaatgcagatgagtgatgacgctatttatctaccaggtgga
 catggtactgtatacgattttgccaaatagagaaattagctgatattttacttcaattt

aaaaatagtaataaaaataatctcttcagtatgtcatggacctagtgcgtttgtagggtga
 aaagatgcaataatcactatctagtagatgggtgtcaaaaataacttcatttactgatagt
 gaagaaaaagcaatgggatttgaaaataaagtaccatttttaactcaatctaaattagaa
 gagcaaggtgcaaattttgtagtgaagatgactttacatctcacgtagaaaaagacggt
 5 caatttatcactggacaaaatccccaatcaagtgaagacattggtaaagcacttgcaaat
 gaattaaaaataa

Sequence 2824

VYIFMSKKVLFVLSTSTSQFTDGTETGLWLEEAGAPYNILTEEGINVDVISIKGGK*NLDP
 10 NSVSNESLNQYAKFVSHLNDTPSIENVNADEYDAIYLPGGHGTVYDFANNEKLADILLQF
 KNSNKIISVCHGPSAFVGVKDANNHYLVGVKITSFTDSEKAMGFENKVPFLTQSKLE
 EQGANFVVKDDFTSHVEKDGQFITGQNPQSSDIGKALANELK*

Sequence 2825

15 Contig_0458_pos_1499_1038
 is similar to (with p-value 4.0e-51)
 >sp:sp|O31408|ARGR_BACST ARGININE REPRESSOR. >gp:gp|Y09546|
 BSARGR_1 B.stearothermophilus argR gene. NID: g2369705.
 gtgttaattgtgccaaaaagtcagtgagacataataaaaaaagagagataatttcaaat
 20 gaacaaatagaaacacaagatgaactagttaaacggttgatgagtgatgatttaaatggt
 acacaagctactgtttcacgagatattaaagaattgcaattaattaaagttcctgcacct
 acagggcaatatgtttatagtttaccaaatgatcgtagatatcatccattagagaagttg
 ggtagatatttaattggttcatgttgaacattgagggtagtgtaattctactagttcct
 aaaacgcttcctggtaattgctcaatccattgggtgctatacttgatcaaattgattgggat
 25 gaggtacttggtacaatttggtggtgatgatacatgcttacttattgtcgcagacgaagaa
 gcgagtgaaagaatcaaaactcgaattttcaatttattataa

Sequence 2826

30 VLIVPKKSVRHIREIISNEQIETQDELVKRLNEYDLNVTQATVSRDIKELQLIAVPAP
 TGQYVYSLPNDRRYHPLEKLGRYLMDSFVNIEGTGNLLVLKTLPGNAQSIGAILDQIDWD
 EVLGTICGDDTCLLICRDEEASEEIKTRIFNLL*

Sequence 2827

35 Contig_0460_pos_2481_3032
 is similar to (with p-value 2.0e-28)
 >sp:sp|P26646|YHDH_ECOLI HYPOTHETICAL 34.7 KD PROTEIN IN MR
 EB-ACCB INTERGENIC REGION (ORF1) (O324). >pir:pir|JS0688|JS0
 688 hypothetical 35K protein (fabE 5' region) - Escherichia
 coli >gp:gp|M80458|ECOACOAC_1 E.coli biotin carboxylase and
 40 biotin carboxyl carrier protein (fabE) and ORF1 35 kDa prote
 in genes, complete cds. NID: g145172. >gp:gp|U18997|ECOUW67_
 183 Escherichia coli K-12 chromosomal region from 67.4 to 76
 .0 minutes. NID: g606010. >gp:gp|AE000404|AE000404_7 Escheri
 chia coli K-12 MG1655 section 294 of 400 of the complete gen
 45 ome. NID: g2367207.
 atgtctattgaaggtaaagaagtgcttgtagcaggtgccactggaggcgtcggaaacgatt
 tcattactcatgttaataaacttagggatgatgttattgcaagtacgggtagagatgac
 gcggaggaaaaacttaaaaagcttggtgctaaagaagtaattggccgtttaccagagat
 aatagtaaacattagagaagagaacatggcaggcagccattgatccagttggtggtgaa
 50 aacttaccgtacatcgtaaacgattggataacaacggaagtgttgcatattggcatg
 actggtggtgaacaattttgaaacaaccgtcttctcttcattttaagaggagcaagtata
 attggtatcgatttacttactccaattaaactaagaaaacgtgtttggagaagactt
 gcaaaaagacttaaaaccacaacaattacatgacatcaaacatgttatttcattcgatgaa
 atcccaaaagccatcgatcaagtcataatcataataataactggacgtattgtcattgat
 55 ttcaatgtttaa

Sequence 2828

MSIEGKEVLVRGATGGVGTISLLMLNNLGYDVIASTGRDDAEKLLKLGAKVIGRLPED
 NSKPLEKRTWQAAIDPVGGENLPYIVKRLDNNGSVALIGMTGGNNFETTVFPFILRGASI

ccttcggaatgtgtggtccttaaagtcattgcaatcgaagattttgaaacaggtaaaaag
actgacttacaaaatgatgaagtcagcgatataactttacctaaagcgaatgtaataaag
atatactttaatgaaggatttattgctttgcgtccttctggtacagagcctaaaattaaa
ctttatgtatcactttctgtgaccattttgacgtagttgcacaaaaaatgaatgatgct
5 atatttaactcttaa

Sequence 2832

MIISYLHNINKLKNNSFLDLTQNKNGNMIYFKKEVILKMIDNWIDVLDSESLVKDFYNNQT
SEEQEGGLDITLSFGTAGIRGKFGLEGRLNKFTVSKVALGFAHYLTSSIAHPVVVIHYD
10 TRHLSPDFAQIIANILASLDIKVYLADTYRTTPDLSFAVRYLQADAGIMITASHNPKDYN
GIKVYGEFGAQLSTDDARSLEYIDKLGHPLHINLPSLTTEQOTLIHVPSEVREDYFKN
VQDLVGITPQSDLKVVFTSLHGTSPVVPDILSSLNFNQFELVASQCEPDSDFSSVVSAN
PEDHKAQDSIELANHIDADLLIGTDPDADRLGIVERDAEGNIHYNGNQIGALLNYRI
KQTEGLPNRIMFQSIIVSGGLAKSLAQYHNVNFKEVLTGFKYIAAEIRHLSPEQNFIFGYE
15 ESYGFLARPFVRDKDAIQIVPLMIKYAAELKNKGRMLKDELEDITRNVGNFNDKLFSTF
EGTQGKAKIENIMTQFRSETPSEMCGLVIAIEDFETGKKTDLQNDDEVSDITLPKANVIK
IYFNEGFIALRPSGTEPKIKLYVSLSCDHFVVAQKMNDIAFNS*

Sequence 2833

20 Contig_0464_pos_2837_3322
is similar to (with p-value 4.0e-24)
>sp:sp|P74561|HIS4_SYNY3 PHOSPHORIBOSYLFORMIMINO-5-AMINOIMI
DAZOLE CARBOXAMIDE RIBOTIDE ISOMERASE (EC 5.3.1.16). >gp:gp|
D90916|D90916_42 Synechocystis sp. PCC6803 complete genome,
25 26/27, 3270710-3418851. NID: g1653715.
gtgaccactaagcctatagaagtggtggtggtggtcgttcaaaacaaacaattgaaaat
tatattcattcaggaatagactattgtattgttaggtacaaaaggtatccaagatatagag
tggttaacacatatgacacatcaatttccaaataaaactctacttatccgtagatgctttt
ggagagaaaaataaagattaatggatggaaagaggatgctaaactcaatttatttgattat
30 gttgccaaaattgagcatttacctttgggtggtggtgatttataccgatatttcgaaagat
gggaacatttcttgacctaattttgattgacaggtcgtctcgcactttatacatcgttg
cctgtaattgcttcaggaggtatttagacatcaagaggacttgtttcgattagaatcggtta
aatgttcatgctgctattgtaggaaaagcagcacatctggatgaattctgggagggtatta
tcttga

Sequence 2834

35 VTTKPIEVGGGIRSKQTIENYIHSGIDYCVGTGKIQDIEWLTHMTHQFPNKLYLSVDFA
GEKIKINGWKEDAKLNLFDYVAKIEHLPLGGVIYTDISKDGLSGPNFDLTGRLALYTSL
PVIASGGIRHQEDLFRLESINVHAAIVGKAAHLDEFWEGLS*

Sequence 2835

40 Contig_0464_pos_4161_4709
is similar to (with p-value 5.0e-40)
>sp:sp|P44434|HIS2_HAEIN PHOSPHORIBOSYL-AMP CYCLOHYDROLASE
45 (EC 3.5.4.19) / PHOSPHORIBOSYL-ATP PYROPHOSPHOHYDROLASE (EC
3.6.1.31). >pir:pir|A64071|A64071 phosphoribosyl-AMP cyclohy
drolase (hisIE) homolog - Haemophilus influenzae (strain Rd
KW20) >gp:gp|U32730|U32730_6 Haemophilus influenzae Rd secti
on 45 of 163 of the complete genome. NID: g3212191.
50 atcaatgaagaagcttatcaaaaaactctgaaagaaaagaaagtaaccttcttctctaga
tctaaacacgctttatggactaaaggtgaaacttctggtcatttccaacacggttgagagt
attcatctagattgtgatcaagatgcaatcttaatacaaaagtgtgccacaaggtcctaca
tgtcacactggaagtctgagttgttttaataagtgaattgaatcgcgctttaaaattcaa
gcattagcacaacgattcatcaaaagtgtctaaaagcaatcaatctaactcttacactcaa
55 tatttattaaaggaaggtatcgagaaaatatccaagaaatttggtgaagaggcatttgaa
gttggtgataggtgcgataaaacataatcgtgaagaagttattaatgaaacagcagatgtc
atgtatcaccttttctgttactacatagtttggtatattccatttccagaagtagaacag
gtactagcgcatacgccatcaaaaaagaaataatttttaagcgagcgcaaaaaggttcaa
gaatggtaa

Sequence 2836

MNEEAYQKTLKEKKVTFFSRSKQRLWTKGETSGHFQHVESIHLDQDAILIKVM:QGPT
 CHTGSLSCFNSEIESRFKIQALAQTIHQSAKSNQSNSTQYLLKEGIEKISKKFGEAEFE
 5 VVIGAIAKHNRREEVINETADVMYHLFVLLHSLDIPFSEVEQVLAHRHQKRNNFKGERKKVQ
 EW*

Sequence 2837

Contig_0465_pos_10122_9604

10 is similar to (with p-value 3.0e-49)

>gp:gp|Y13052|SSK3MECA1_3 S.sciuri mecA1 gene, strain K3(MM
 2). NID: g2791901.

gtgagcatatatattatgaatatatgagagtcattttatactataatcaatcgcgacacga
 ttgaaaagaggtggagatgtgaaacaagaacaaatgaggttagcgaatcagctttgtttt
 15 tcagcatataatgtaagtcgtttatttgctcaattttatgagaaaaagttaaaacagttt
 ggtataacttattctcagtatattagtagtactgacgttatgggaagagaatcctcaaaaca
 ttaaattcaattggtagacatttggtattatctagtaatactttaaccccttactaaaa
 agacttgagcaatctggctgggttaaaagagaacgtcaacaatctgataaacgacagttg
 ataattacgttaactgacaatgggcaacaacaagaagctgttttgaagcaatttca
 20 agttgcttaccacaagaatttgatacagactgagtagatgaaacgaaatatgtgtttgaa
 gaactagagcaaacattaaaacatctcatagaaaaataa

Sequence 2838

VSIYIMNICESFYTIINRRLKRGDVKQEQMRLANQLCFSAYNVSRLFAQFYEKKLKQF
 25 GITYSQYLVLTLWEENPQTLNSIGRHLDLSSNTLTPLLKRLEQSGWVKRERQQSDKRQL
 IITLTDNGQQQAEAVFEAIISSCLPQEFDTTEYDETKYVFEELEQTLKHLIEK*

Sequence 2839

Contig_0465_pos_9362_8244

30 >sp:sp|P47169|YJ9F_YEAST HYPOTHETICAL 161.2 KD PROTEIN IN N
 MD5-HOM6 INTERGENIC REGION. >pir:pir|S57160|S57160 sulfite r
 eductase homolog YJR137c - yeast (Saccharomyces cerevisiae)
 >gp:gp|Z49637|SCYJR137C_1 S.cerevisiae chromosome X reading
 frame ORF YJR137c. NID: g1015875.

gtgaattggaagaatttaaatgcattacaaagaacaagtcattaatcctatgtctgaaacc
 ctcacatcgatgtttgaacaacaggaattgatgtaatcatgggaaaggtaaacttgta
 gatgctcatacaatagaggttaataatacaactttacaatcagattatattgttatagca
 actggacaacatagtcattcaattagatattgagggtaagaatatacgcagtagatgcg
 gaatttttatcaatgcaatccttaccggatagtagtacttttattggagcaggtattatc
 40 agtattgaattcgcttctatcatgatcaaatcaggtgtagaggttaattgtggttcacat
 acaaatcatgcacttgaagggtttaacgaatcacacgtcaataaatttaattcaaaagtta
 aaagtgaagggtgtaaaattttacttttagtgagaataccaagtcagttaaaccgaatgca
 caacgttttatagtagaaactgagtcctggaaagatgattgaaacagattatgtactggat
 gcaaccggtagaaagcctaattgttcagcaaatagggttgaaaaagtggtatactattt
 45 agtgatagaggtattgaggttgacgattatttaagaacaaatgtgaaaaatatatacgca
 agtggggacgttatcaataaaatgattcctaaacttactcctacagctacatttgagtct
 aattatatcgctgcccatatccttggttgattgaatacagatgccattcagtatccaccaata
 ccttcagtgcctttattcattgcctcgctttatctcaaatagggtgtcacagttagcgaggct
 aagaaagatgatacgtatatgattaaagatataccattcggaagacaaatggtatttgag
 50 tatcaaaacgaaacagaggctgaaatgtcaattgtattagatagtcacaaacgtttagta
 ggagcagagatttatggttaatgacgctggtgatttggttaattccttagtctttatcatt
 aatcaaaaacttactgcacaagacttaataaaaaatatttttgcatttctcggagcttct
 agtggcgcttatagatttattgaaattggcaatgatgtag

Sequence 2840

VNWKNLMHYKEQVINPMSETLTSMFEQQGIDVIMGKGLVDAHTIEVNNTTLQSDYIVIA
 TGQSHSHQLDIEGKEYTHDSREFLSMQSLPDSITFIGAGIISIEFASIMIKSGVEVNVVHH
 TNHALEGFNESHVNKLIQKLKDEGVKFYFSENTKSVKPNQRFIVETESGKMIETDYVLD
 ATGRKPNVQQIGLEKVGILFSDRGIEVDDYLRTNVKNIYASGDVINKMIPKLTPTATFES

NYIAAHILGLNTDAIQYPPIPSVLYSLPRLSQIGVTVSEAKKDDTYMIKDIPFGRQMVFE
YQNETEAEMSIVLDSHKRLVGAEIYGNDAGDLVNLVFIINQKLTAQDLNKNIFAFPGAS
SGVIDLLKLAMM*

5 Sequence 2841

Contig_0465_pos_8159_7755

>pir:pir|A34231|A34231 sulfite reductase (NADPH) (EC 1.8.1.2) - Salmonella typhimurium >gp:gp|M23007|STYCYSJIHA_1 S.typhimurium NADPH-sulfite reductase flavoprotein component (cysJ), NADPH-sulfite reductase hemoprotein component (cysI), and 3' phosphoadenosine 5'-phosphosulfate sulfotransferase (cysH) genes, complete cds. NID: g153928.

gtgctcattatcgtaaaactgcaacgaattgtactctgaatgatcaatttaataaatta
gaaacgcattataaaaaatatcataagtatggctctgaaattttgagcttcccttgcaat
15 gattttaataatcaggagccagaattaatcaaagatatttatcgagtatataaatataag
tttggtattactttcccatccatgctaagattaatgttaatggagagcatgaacaccct
ttgtacacattattaaaaatgtaaacaccaggattatttggttcgcaaattaaatggaat
tttactaaattttagtagatcaacagggaatattgttaaacgatttttaccttgtgat
aatcctaaccaaatgaaaaattaataagacaattattaaaaataa

20

Sequence 2842

VLIIIVNTATNCTLNDQFNKLEMLYKKYHKYGLEILSFPCNDFNNQPELIKDIYRVYKYK
FGITFPIHAKINVNGEHEHPLYTLLKCKQPGFLGFSQIKWNFTKFVVDQOGNIVKRLPCD
NPNQMEKLIRQLLK*

25

Sequence 2843

Contig_0465_pos_6454_4568

is similar to (with p-value 1.0e-30)

>sp:sp|P52035|BSAA_BACSU GLUTATHIONE PEROXIDASE HOMOLOG BSA
A. >gp:gp|L77246|BACYACA_17 Bacillus subtilis (YAC10-9 clone
30) DNA region between the serA and kdg loci. NID: g1256615. >
gp:gp|Z99115|BSUB0012_132 Bacillus subtilis complete genome
(section 12 of 21): from 2195541 to 2409220. NID: g2634478.

atgacaaatatcggaataacaaagtataagaggtggataagattgaacttaagtgttacc
35 aatagtccttttacggaagggaagctaaacaaatcaatgaattgcttcacactttatca
tctaaccaacaagtatggttaagtggctatctgatggcaaatcaacaatcaaatacatct
acagattctgtagaacaacataactcagatgacaatacagaagcgatgctacatgaaaaag
gaaccttcagttgagccggaagctagatcgataacaatttttatatggatctgagtctggt
aatgcgcaaggacttgagaaatatttgaaacacggtttatctgatattggaatgacggt
40 acgctcaaatcaatggatgacttttaacctaagaatttgaaagaaagtagaagatttattt
attatcacatctacacatggcgaaggagatccaccagataacgctgttgattacatgaa
tacctccacggacgtaaaagcgcaaaaattggatggggtgagattttcagtattagcatta
ggagatcaaacctatgaattcttctgtcaaaactggtaagattttgataaccggttggct
gaattaggggcagaaaagactttatcatcgtactgattgcgatgttgattatgaagaagac
45 gctgaaaagtggatggccaatgttattaatacagattgattcaacaccggctggtactgaa
agtgaacaagttgttagcgaatctataaaatcagcaaaaagaaagaaatatagtaagtct
aatccttacgatgctgaagtattgaccaatattaatttgatggtagaggttcagacaaa
gagacgagacatataggattattacttgataattttggtgaggaatatgaaccgggagat
tgtgtagttgtcttgcccaaatgatccagctatcgtagacttacttattagcacatta
50 ggttggagtcagaaaacaaagttttaattaacgaggatggagatactttaaatcttgaa
gaggcattaacatcgcattttgaaattactaaattaacaaaaccgtaataagaaaatgct
gcgatatttttgataatgaagagctttctgaaaaaattcaagataaagaatggattcaa
aacacg*tgagggaagggtatttgattgacttattaaatgacttcgcaacgacagaaacta
caacctgaaaatttacatcaattattaagaaagttaccacctagagagtactcaacatct
agtagttataaagcaacaccagatgaagttcacattactgttgagcagtttagatatcaa
55 gcacacggctcggaacggttcagggtgtttgttcagtacaatttgcagagagaatacaagag
ggcgatacaatttcctatctatttaaaacgaaatccgaattttaagtttccgcaagatgaa
tcaacacctgtgattatgataggtcctgggacaggtgttgaccggttagatcctatatg
caagaacgagaggaactaggttttgaaaggaaatacatggttattctttggagatcaacac

Sequence 2844

10 MTNIGITKYKRWIRLNL SVTNSPFTGEQAKQINELLHTLSSNQQVWLSGYLMANQQSNTS
TDSVEQHNSDDNTEAMLHEKEPSVEPEARSITILYQSESNGAQLAEIFEQRLSDIGNDV
TLKSMDDFPKPNLKKVEDLFIITSTHGEGDPPDNAVELHEYIHGRKAPKLDGRVFSVLAL
GDQTYEFFCQTGKDFDNRLAELGAERLYHRTDCDVDYEDAEKWMANVINTIDSTPAGTE
SEQVVSSEKISAKEKKYSKSNPYDAEVLTNINLNGRSDKETRHIELLNDFGEEYEPGD
15 CVVVLPQNDPAIVDLLISTLGWSPETQVLINEDGDTLNLEEALTSHEFITKLTKPLIENA
AIFFDNEELSEKIQDKEWIQNYVEGRDLIDLNDNFATTELQPENLHQLLRKLPFPREYSIS
SSYKATPDEVHITVGAVRYQAHGRERSGVCSVQFAERI QEGDTIPIYLKRNPNFKFPQDE
STPVMIGPGTGVAPFRSYMQEREELGFEGNTWLFFGDQHFTTDFLYQTEWQEWLEDGTL
SKLDVAFSRDTHKKVYVQHKIAENSEQFNRIENGATIYVCGDESKMAKD VHQAIKNVLI
KEQNLSETDAEEYLKOMKRD KRYORDVY*

55 Sequence 2846
MVNTNNHISEELDKNLDEMEFLKANSDFLRGTIEQSLANPITGSITQDDAKLLKFHGSYM
QDDRDLRDERRKQKLEPAYSFMIRVRVPGGKATPEQWIAMDDISQVYANHTTKLTTRQAF
QFHGILKRNKLGQSMKNNHVLDSIAACGDVNRNTMCNPNPYQSQVKEINDYATRISNH
LLPRTNAYHEIWLDEKVLDSSEEEKPIYGNTYLPKRFKIGIAVPPSNDIDVYSODIGLI

AIVEQDELIGFNVITIGGGMGMTHGNTETYPQLGRLIGFIPKEKVVDVCEKILTIQRDYG
 RENRKNARFKYTVDRLGSETWVTEELNRRRLGWEIKAPRDFEFHNGDRLGWIEGINNWNFT
 LFIQNGRVKDTEDYLLKTTLREIAEIHGTGDFRLSPNQNLVIANVSPEKKEEQAIIDKYK
 LTDGKNYTGRLRRNSMACVAFPTCGLAMAESERYLPSLITKIEDLLDESLKEEEITIRMT
 5 GCPNGCARPALAEIAFIGKAPGKYNMYLGGSFKGERLNKIYKENIDENEILESRLPLLLR
 YSKERLDGEHFGDFVIRGVIKVDHGRDFHS*

Sequence 2847

Contig_0467_pos_7413_6943

10 is similar to (with p-value 2.0e-17)

>sp:sp|P45637|YPRA_CORGL HYPOTHETICAL 33.0 KD PROTEIN IN PR
 OB-PROA INTERGENIC REGION. >gp:gp|U31230|CGU31230_3 Coryneba
 cterium glutamicum Obg protein homolog gene, partial cds, ga
 mma glutamyl kinase (proB) gene, complete cds, and (unkdh) g
 15 ene, complete cds. NID: g950194.

atcaaaagtaataggggttagtaagtcagaaaaaatgttgacaatttgatgaagtatat
 acaartgaagaggttagatgatgttattgaaaaggcaaatattattgttaatgcattacct
 gaaacagaagaacaatttacttattaaaaagaaaagatttcatacaaatggacaataat
 20 gcttattttataaatgttaggaaggaacaattgttgatgaggaagtgtcatcaatgta
 ctcaaagatcgattaatcagaatgcttatttagatgttttgaaaaagaaccacttagt
 aaggacaatcctttatatgatattagataatgtgaccataactgctcatattacaggtaat
 gattctaataataatagagaagctacggacattttcaaaaagaatcttgagcattttctc
 aataattatgatgtaattgagaataaagtagacttagattatggttactaa

25 Sequence 2848

MKVIGVSKSGKNVEQFDEVYTIEELDDVIEKANIIVNALPETEETIYLLKRKDFIQMDNN
 ALFINVGRGTIVDEEVLINVLKDRILRHAYLDVFEKEPLSKDNPLYOLDNVITITAHITGN
 DSNNNREATDIFKKNLEHFLNNYDVIEKNKVDLDYGY*

30 Sequence 2849

Contig_0468_pos_13714_9209

>sp:sp|P39812|GLTB_BACSU GLUTAMATE SYNTHASE [NADPH] LARGE C
 HAIN (EC 1.4.1.13) (NADPH-GOGAT). >gp:gp|Z99113|BSUB0010_138
 Bacillus subtilis complete genome (section 10 of 21): from
 1781201 to 2014980. NID: g2634090. >gp:gp|Z99114|BSUB0011_9
 35 Bacillus subtilis complete genome (section 11 of 21): from 2
 000171 to 2207900. NID: g2634230.

gtgtgcattcatgacaatgagaaattaaaaaggactatacagattatcgtgaagagcat
 gatgcgtgtggtatttgattttatgccaatatggataataaaagatctcacgatattata
 40 gaaaaactctttagaaatgttaagacggttagatcatcgtggtggagtaggtgccgatggt
 attactggtgatggtgcaggaattatgacggagataccataccaattgttcgaacaatta
 acagaattcaaagtccccggcgaaggatattatgccgtgggattattttttctaaagag
 aaagtttagagattcaattcacgaagagatgtttaatcaatattttgaaagtgaaggttt
 aaagtcattggatatagagatgtgccagtagatactcgcgctattgctcaacatgttgca
 45 gatactatgccttatattcaacaagtattgttgacatcacaggtgtaaaagaagttgaa
 aaacgattgttttttagcaagaaagcaaatgaaaaatatagtgaacacaatccatagat
 ttatattttacaagtctctctcatagaacgattgtttataaagggttggttacgttcggat
 caaattaaaggcttatatttagacctacaaaatgaggcatatcaatcaaaattaggactt
 gtacactccccgcttttagtactaatacatttccaagttggaaacgtgcacatcccaatcgc
 50 atgcttatgcacaatggtgaaattaataaccattaaaggttaacgtaaaactggatgcgagca
 cgccaaaataaaactagttgaaacattatttgaagatgagaaagataaaggtgcattttatt
 gttgatgaagatggttagtgactcatcaatagttgataatgcgttagagttcttatcatta
 gcaatggagcctgaaaaagcagcgatgttattattccagagccatggttatacaatgaa
 tctaacgataaaaaagttcgctctttctatgaattttatagttatttaattggagccatgg
 55 gatggaccaactatgatttccttttgcaatggagataagataggtgcattgactgataga
 aatgcatgaaagacctgggctgtatatacaataactaaagacaattttattgtttttttcc
 gaagtaggtgtcattgatgtccagaagaaaatgtagcatttaaaaggacaacttautcc
 ggaaagttattactttagactttttgcaaaaataaggtttagaaaaataatgagctaaaa
 actaatattgctaattgagttgccctacgaacaatggctaaaagattataaaaaataaaat

gatttagataatatttattaccaatcttccgactgggatgatcaaacactcttccgctta
 cagaaacaatttgcttacactaaagaagatatcaataaatatgatgacagatttagtcatc
 aataaaaaagatcccatcggagcgatgggatatgatgcacctattgcagttcttaatgat
 aagcctgagtcactatttatttttaacaattatttgcacaagtaccaatccaccc
 5 atcgatgcttatcgagaaaagattgtaactagtgaactttcataatttaggttcagaagg
 aacctcttatgtcctgatgaatcagtttttagaaagaattcaattaaaaaaccagtttta
 aatgaagcgcaattatcatcaatcgatcattcgtattttaatgtaacgtatttatctaca
 ctttatcacaggtgatttggaagtagcttaaatgaactaggaaccgagcaatacaggt
 gtacutgaaggtgcgaaaattttggtgttagacgatacgtctttaactcacgaaaatagt
 10 tatgcaatgccaattttattagcgttaagtcacgtgcatcaattattaattcgagaagga
 ttaagaatgagaccagcctcattgcgcagtcgggtgaaacacgagaagttcatcacgtt
 gcatgtttacttgggtatggtgcaaacgctgtagttccatatttagcgcaacgaacgatt
 gaacaattaacgcgtcaaggtcaactttcaggaactgtcgctgaaaatgttgctacgtat
 accaatgtattgtcagaaggcgttattaaagtgtggctaaaatggcatttctactgta
 15 caaagttatcaaggagcacagatatattgaagcggtaggttatcgaaatagcgtcattgaa
 aatattttacaggtacacagtcacaaattatctggtataagttatgaacaaatagacaaa
 gagaataaagcgagacaaagtgtgattctgattatcttgaatccggaagtgtattccaa
 tggagacagcaaggtcagcctcattgcgcagtcgggtgaaacacgagaagttcatcacgtt
 gcatgttagagaaaatgattacgagttatttataaaatctccaaaactgtaaatttaaaa
 20 cgtacggatcatattagacatttattagaattcaagacacgccaatctattgatattagt
 cgtgttgaaaccagcaagtgaattcgtaaaacgttttaatacaggagcaatgagttacggc
 tctatctcagcagaggcacatgagacgttggctcaagctatgaatcaaatggaggtaaa
 agtaatagtggagaaggtggagaagattcttcacgttacgaaattcaaaaggatggaagt
 aataagataagtcgattaaagcaagttgcatcaggtcgttttgggtgacgagtgattac
 25 ttgcaacatgcaaaagaaattcaaatataagtcgcacaaggcgttaaacagggggaaggt
 ggacaactaccaggttcaaaagtatatccatggattgctgagactagaggttcgacacca
 ggtataggatttaatttcaccaccaccacaccatgatatttattcaattgaggacttagca
 cagctcatcattgatatttataaaatgcaaatagaagagctgatattgcagttaagc":cgta
 tcaaaaactggcgttggaactatagcttcaggggtagctaaagctttcgccgataaaaatt
 30 gttataagtggttatgatggaggtacaggtgcatcgccataaaacaagtattcaacatgca
 ggtttgcatggagataggccttgccgaaacacatcaaacacttaaatataatgatttg
 cgtagtcgcgtataaattagaacggatggttaagttactgacgggttaagatgtagcttat
 gcttgtgcgcttggtgcagaagaatttggtttcgcaacagcaccacttgttgttttggg
 tgtattatgatgaggggttgtcataacgatacgtgtccagtaggggttgcaacacaaaac
 35 aaagatttaagagctttgttttagaggttaaggcacagcatgtagtttaactttatgtattt
 atagctgaagaattacgtgaaattttggcttcacttgggttagaaacagtagaagagtt
 tagggaagaacagatcttcttcaacgttcgacgcaattgaaacaaatagtaaaagcagct
 tcgcttcaaatagaacgtttaatagaacaatttgacgggttaatacgaagagatatca
 caaaaccatcatcttgatgaaggatttcgatttgaattatctgtaccagacgcacgctat
 40 agtattgaaaacgggcactctttaccggaaattatgttgttaataatgaacagcgagat
 gtagggtgaattacaggtagtgcgatagctaaacaatatggagaagaaggattacctgaa
 gatagataacttgcttacctgaaggtcatgcaggtcaaagcttagctgcatatgcacca
 cgcggttaacaatccatcataccggtgatgctaatgactacgtaggtaaaaggattgtcc
 ggtggaactgtcatcgtaaatgctccaaatagtcaacgtgaaaatgaaattatagcagga
 45 aatataaaacttttacggggtctctagaggtaaagcgtttatcaatggtaaaagctggtag
 cgtttctgtatcagaaatagtgggtgcagatgtttagtagaaggatttggtagcatgga
 cttgaatatatgacagggggacatgtcattatcttaggagatgttggaagaactttggc
 caagcatalgagcggggcgtaagttatatttctctctgacgtggagaaatttaaaag
 gttaatgcgcttgaaactttagaattcagtagcatacgttttgatgaggaaaaatctctt
 50 atcaaagacatgcttgaagcacattttaagcatcacgtagtaacaaagcacgccaatta
 cttagcaaatttgacaatattgaaaagttagcaattaaagttattccgaaagattacaaa
 ttaatgatgcaaaaatttgatttgaaaaaacgtcaaatggaacgtgaagatgaagcaaca
 ctggcagcgttttatgatgacagagaaacaattgaacaagagctacagccagcagtcatt
 tatta

55

Sequence 2850

VCIMYNEKLKKGLDYREEHDACGIGFYANMDNKRSHDIEKSLEMLRRLDHRGGVGADG
 ITGDGAGIMTEIPYQLFEQLTEFKVPEGYYAVGLFFSKEKVRDSIHEEMFNQYFESEGF
 KVIgyrdvpvdtraiaqhvadtmppyiqqvfditgvkevekrflarkqiekyssetqsid

LYFTSLSHRTIVYKGLRSDQIKGLYLDLQNEAYQSKLGLVHSRSTNTFPSWKRAHPNR
 MLMHNGEINTIKGNVNWMRARQNKLVETLFEDEKDKVHFIVDEDESDSSIVDNALEFLSL
 AMEPEKAAMLLIPEPWLYNESNDKKVRSFYEFYSYLMEPWDGPTMISFCNGDKIGALTDR
 5 NGLRPGRYTITKDNFIVFSSEVGVIDVPEENVAFKQQLNPGKLLLVDFLQNKVVENNELK
 TNIANELPYEQWLKDYKNKNDLDNIYYQSSDWDQTLFRLQKQFAYTKEDINKYMTDLVI
 NKKDPIGAMGYDAPIAVLNDKPESLFNYFKQLFAQVTNPPIDAYREKIVTSELSYLGSEG
 NLLCPDESVLIERIQLKKPVLNEAQLSSIDHSYFNVTYLSTLYTGDLESSLNELGNRAIQA
 VHEGAKILVDDTSLTHENSAMPILLALSHVHQLLIREGLRMETSLIAQSGETREVVHV
 10 ACLLGYGANAVVPYLAQRTIEQLTRQGQLSGTVAENVATYTNVLSEGVKVMKMGISTV
 QSYQGAQIFEAAGLSNSVIEKYFTGTQSKLSGISIEQIDKENKARQSDSDYLESGSVFQ
 WRQGGQHAFNPRTIFLLQHACRENDYELFKKFSKTVNLKRTDHIRHLEFKTRQSIDIS
 RVEPASEIVKRFNTGAMSYGSISAEAHETLAQAMNQIGGKSNSGEGGEDSSRYEIQKDG
 NKSIAIKQVAGSRFGVTSYDLQHAKEIQIKVAQGAKEPGEQQQLPGSKVYPWIAETRGSTP
 GIGLISPPPHHDIYSIEDLAQLIHDKNANRRADIKLVSKTGVGTIASGVAKAFADKI
 15 VISGYDGGTGASPKTSIQHAGLPWEIGLAETHQTLKLNDRSRVKLETGKLLTGKDVAY
 ACALGAEFEFGATAPLVVLGCIMMRVCHNDTCVGVATQNKDLRALFRGKAQHVVNFMYF
 IAEELREILASLGLTVEELVGRDQLRSTQLKPNSKAASLQIERLIEQFDGVNTKEIS
 QNHHLDEGFDLNYLYPDARYSIENGHSFTGNVYVNEQRDVGVITGSAIAKQYGEGLPE
 DTILAYTEGHAGQSLAAYAPRGLTIHHTGDANDYVVGKLSGGTVIVNAPNSQRENEIAG
 20 NINF*GASRGKAFINGKAGERFCIRNSGADVVEGIGDHGLEXYMTGGHVIILGDVCKNFG
 QGMSGGVSYIFSSDVEKFKKVNALLETLEFSSIRFDEEKSLIKDMLEAHFKHTRSNAKARQL
 LDQFDNIEKLAIKVIKDYKLMMQKIDLKKRQMEREDATLAIFYDDRETIEQELQPAVI
 Y*

25 Sequence 2851
 Contig_0468_pos_9170_7728
 is similar to (with p-value 5.0e-25)
 >sp:sp|P33019|YEH_ECOLI HYPOTHETICAL 36.9 KD PROTEIN IN LY
 SP-NFO INTERGENIC REGION. >gp:gp|U00007|ECOHU47_49 47 to 48
 30 centisome region of E.coli K12 BHB2600. NID: g453983. >gp:gp
 |AE000305|AE000305_3 Escherichia coli K-12 MG1655 section 19
 5 of 400 of the complete genome. NID: g1788479.
 atgaaatatgataaacagtcgctatcagaattgtctttggttagaccgtctttcgaaatcat
 gaagcgtttcaacaacgcttcactaaagaagatgcttcgattcagggcgcgctgtatg
 35 gattgtggaacacctttttgtcaaaactgggcaatcttatggaagagaaacaataggatgc
 cctatttggaattatatactgagtggaacgacttagtctatcatcaagattttaaagct
 gcttacgaaagattgagagagacgaataattttcctgaatttacaggaagagtttgcct
 gcaccatgtgagcaatcatgtgttatgaaaaataatagagaatccgtggcgattaaaggt
 at: gacgtacaattattgatgaagcatatgagaatgagtggttcacccgcattccct
 40 gaagatcataaagaccaacgagttgctatcgtaggtagtggtccagcgggacttacagca
 gctgaagaattaaactttaaaggctataaagttactgtttatgaaaaggcgcatgaacca
 ggcggcttgctaattgtatggtataccaaatatgaaactagataaagacgtaatacgtcga
 cgtgtatcacttatgaaagatgctggggttttatttaaacaggcggttgaaattggcgctc
 gatgtgagccgtgaaacacttgaaagaaaattatgatgctattattttatgcacagggtgct
 45 caaaatgcgagagatttaccattggaaggacgaatgggctctggtattcattttgcaatg
 gactatcttactgaacaaacacagtatctaaatgggtgagattgaaagtttgagcattact
 gctaaagataagaatgtaattatttataggtgctgggtgatactggtgcagactgtgtagcg
 acagcattacgtgaaaactgtaaatctattgttcaatttaataaatatacgaacagcct
 gaagagattacttttgaaagtaatacttcctggccattagcaatgcctgttttcaaaatg
 50 gattatgcgcataaggaatatgaagctaaatttggtcaagaaccaagagcctatggtgta
 caaacaatgcgctatgatgttgacgagttaggaaatgttaaaaggcttatatacacaata
 ttaaaagaaacgcctgatggcatggtgatggaagatggaccagaacgattttggccggct
 gatttagtcttattatctatagggtttgggtgactgaaaccactgttccgcatgcgttt
 gatatacacaccgagcgttaataaattgtagctaataatgatacaaatatcaaaactaatcac
 55 gctaaaatatatttgctgcaggagatgcaagacgaggtcagagtttgggtgtttgggcaata
 aaagaaggctcgtagagtagcacattctgttgatcaataacttaagtaagaagttctagtgt
 taa

Sequence 2852

5 MKYDKQSLSELSLVDRLSNHEAFQQRFTKEDASIQGARCMDCGTPFCQTGQSYGRETIGC
PIGNYIPEWNDLVYHQDFKAAAYERLRRETNNFPEFTGRVCPAPCEQSCVMKINRESVAIKG
IERTIIDEAYENEWVHPAYPEDHKDQQRVAIVGSGPAGLTAEEELNFKGKYKVTVYEKAHEP
GGLLMYGINPMKLDKDVIRRRVLTMDKAGVLFTKGVEIGDVSRRETEENYDAIILCTGA
QNRDLPLEGRMGSGIHFAMDYLTETQTYLNGEIESLSTAKDKNVIIEAGDTGKADCVTA
TALRENCKSIVQFNKYTKQPEEITFESNTSWPLAMPVFKMDYAHKEYEAKFGQEPRAYGV
QTMRYDVDELGNVKGLYTQILKETPDGMVMEDGPERFWPADLVLLSIGFVGTETTVPHAF
DIHTERNKIVANDTNYQTNHAKIFAAGDARRGQSLVVVAIKEGREVAHSVQYLSKEVLV
+

10

Sequence 2853

Contig 0468 pos 6592 5564

>gp|D85230|PEEGLTD_4 Plectonema boryanum URF141, ORF243,
NADH-dependent glutamate synthase large subunit (gltB) and
15 small subunit (gltD) and URF289 genes, partial and complete
cds. NID: ql339947.

gtgctaagatttcatatcgagggtgaaagtcacatatgaaatcaataacgcaggcttcattt
atgaaagggtattatgtttacatttacgattgcaataatcagttatatattagctaaattt
cctattttacatacgattggggcggttagctattgccatcatttttgcatgatataccgc
20 caagtcatagggtatcctgagcatattcgcccagggtattacggttgcatcgaaacggtta
ttaaatttgcgattatcttatactgggttaaaataaataatgggagatattctaggttaa
ggttggaattactactattatgatattatcgtaattcttcttcaataagtttaacttta
cttttgaatcaaattattaaaggaaataaagatatctctatactacttggtattggtaca
ggagtatgtggagctgcagctattgcagctacagcaccaatctaaaatctaagaaaaa
25 gacattgcaataagtgtaggtattattgcactagttggaactatatttgcacttatttat
acagctatcgaggctatttttaacatacctactataacttatggtgcttgacaggtatc
agtctacatgaaatcgctcaagtcggttttagcagcagggtattggcggttcggaggcaatg
acatttgctttacttggaaaattaggccgtgtggttttacttattccataagtattgtc
ttgattttgtatatgcgttataagtcacactcaagtcgaagtacaacaaaaaatcgatatt
30 ccttacttttcaattggatttattataatggcttgatcaatacatttgttccatkcct
tcattacattatgaatattataaatgttatacaacggtattgtatgttaatggcgatgggt
gctctaggattgaatatcgttttaaagaaggtattttcaaaagcacttaaccatttcatt
gtgatctgtataaactcaatttgcctgctcgtggtgtgactctcttagttacgtctataatg
tttaataaa

35

Sequence 2854

40 VLRFHIEVKVHMKSIQTQASFMKGIMFTTIAIISYILAKFPI LHTIGALAI AII FAMIYR
QVIGYPEHIRPGITFASKRLLKF AII LYLGLKLNMGDILGKGWKL LLDIIVII FSI SLTL
LLNQIIKGNKDII SLLGIGTGVC GAAATAATAPILKSKEKDIAISVGII ALVGTIFALIY
TAIEAIFNIPTITYGAWTGISLHEIAQVVL AAGIGGSEAPT FALLGKLG RVFLLIPLSIV
LILYMRYSKSHSSVQ QKIDIPYFLIGTIFIMACINTFSEMI PSLLMNII INVITTLCLMLMAMV
LGLNIVLKEISVSKALKPEIVICITISICLSGVTLI VTSIMFK*

Sequence 2855

45 Contig 0469 pos 5825 5232

is similar to (with p-value 2.0e-40)

```
>gp:gp|Z677739|SPPARCETP_2 S.pneumoniae parC, parE and transposase genes and unknown_orf. NID: q1490398.
```

atgtaaactcttgagttatctgattgggtgcattcccaagcgggttaattattggtaaat
50 ttttttaaaaaagatataagacaatacggtagtggaatactggagcaactaacagt
cgtgttcttggaagaccagctggatttatagttacgttttagatatttcaagggatt
attacagtcttttccactatgggtccagttcatgaggatgggtgtataagcacctt
tttacaatggtttaatagtaggattgttgcatactcggtcacgtgtatccaatat
ctgaaatttaatggcggaagcagtagctaccagtgaggagttgtattaggtgtcaat
55 cctattttacttcttatcttggaattatcttttttagtgtattaaaaatctttaatat
gtttctttatcaagtatcattgcagcaattagttgtgtgattggttcaatcatcattcat
gattatattttacttgcgtgttagcggaattgtttcaatcatattaaatctgcacaaa
tctaatatagttagaatttttaagggaagaacctaataatggaatgtaa

Sequence 2856

MLILSYLIGAFPSGLIIGKLFFKKDIRQYGSNGTGATNSFRVLGRPAGFIVTFLDLFGKF
ITVFFPIWFPVHADGVISTFFTNGLIVGLFAILGHVYPIYLKFNGGKAVATSAGV.LGVN
PILLILAIIFFSVLKIFKYVSLSSIIAAISCVIGSIIHNDYILLAVSGIVSIILIRHK
5 SNIVRIFKGEPPKIKWM*

Sequence 2857

Contig_0470_pos_7843_9180

is similar to (with p-value 3.0e-97)

10 >gp:gp|Y13052|SSK3MECA1_1 S.sciuri mecA1 gene, strain K3(MM
2). NID: g2791901.

atgagtaaggttaatcatttgatagttgaagatgaacggtattttgcacattctggacgg
attaaatattatccactagttattgatcatggttatggagccacgttaatcgatgtagac
ggtaagtccttatattgatttattagcaagtgcaagttcgcaaaatgtgggtcagcgtccg
15 aagccagtagtcgaagcaattagaaccaaactgagaaattcattcattatacaccagca
tatatgtatcatgaaccgttagtgcgattatcaaaaaaattatgtgacattgctcctgga
aattatgagaagagagttactttcggattaagtggctctgatgctaattgatgggattata
aagtttgcacgagcgtatacaggacgtccatatacataagttttactaatgcataccat
ggttcaacatttggttcattatcgatgtcttcaattagtttgaatatgctgaagcattac
20 gggccattacttaattgattttaccatataccttttcccataagtatagggggatgttt
gagcaagctaaagcctaacacggttgaggaaatatttagctcctttaaagaaatgtttgca
aaatacgtccctgcggaggaagttgcatgtattgtagttgaaacaattcaaggtgacggt
ggcttacttgaaccggtaccaggttattttgaagcattacaagagctttgccacgtcac
aatattcttattgcagttgatgatatacaacaaggattaggtcgtagaggaaagtgaggt
25 tccgtagatcattattatttactccagatttaattgacattttgaaagtcatttagctgga
ggtttaccaatgtctgcgattgtaggtcgtaaaagaaatcatggaagtcctgaagcacct
gctcatttatttacaactggtgcaaatcctgtaagttgcgaagcagccttagcaacgata
aagatgattgaagatgaagatttactaaacgcttcatggaaaaagggagttacgttaga
aaaagaatagaccatggatagaacggttatcaatatgtaggtgatgttcgaggtattgga
30 ttatcgattggaatagacatagtatcaaataaaattgagaaaactagagattctgaagca
gcattaaagatatgtaattactgctttgaaaatgggtgtgattatcatagcagttgcgggt
aatgttttaagatttcaaccaccacttgtaattacctataagcaacttgataaagcatta
gatacaatagaacaggcgcttgaaaagttggaagaggagaattaaatcaatatgacatt
agtgggtcaaggttggttaa

35

Sequence 2858

MSKVNHLIVEDERYFAHSGRIKYYPLVIDHGYGATLIDVDGKSYIDLLASASSQNVGHAP
KPVVEAIKNQTEKFIHYTPAYMYHEPLVRLSKKLCDIAPGNYEKRVTFGLSGSDANDGII
KFARAYTGRPYIISFTNAYHGSTFGSLSMSSISLNMRRKHYPGLNGFYHIPFPDKYRGMF
40 EQAKPNVVEEYLAPLKEMFAKYVPAEEVACIVVETIQDGGLEPVPVPGYFEALQELCHAH
NII.IAVDDIQQGLGRGTGWSSVDHYFTPDLMFTGKSLAGGLPMSAIVGRKEIMEST.EAP
AHLFTTGANPNVSCEAALATIKMIEDEDLLNASWKKGSYVRKRIDPWIERQYVGDVVRGIG
LSIGIDIVSNKIEKTRDSEALKICNYCFENGVIIVAVAGNVLRFPPLVITYKQLDKAL
DTIEQALEKLERGELNQYDISGQGW*

45

Sequence 2859

Contig_0471_pos_8235_6703

>gp:gp|Z46863|ACRBDOXN_10 Acinetobacter sp. cysD, cobQ, sod
M, lysS, rubA, rubB, estB, oxyR, ppk, mtgA, ORF2 and ORF3 ge
50 nes. NID: g2462044.

atggaagaagaattggtttgatagacattggttccaacacgattcgacttggttatattt
ggctacaataaaaaaactgggtcaatgaaatactgaatataaaaacacctgcacgttta
agtcaatatctcactaagccaatgaaatgaatgatgaaggtattcatgttttaaaagag
actttaagcagtttttagaaaagttgcggataaaattaacgttgatgcattatatcccatc
55 gcaacagctgctacgtcaatccttaaaaaatcgtgaagctatcattaaagaaattaaacaa
gatattcatatcgaaattcaaatgtacctgaagaagatgaagcattttacggttactat
gcgattacacatactactgatattgaaaatggaatttctgtcgatatcggtggcggttct
accgaagttacccttttcaaagacaaacaactaaagaggctcatagctttccatttggc
gtggtatcacttaagcgtcagttttttggtgataaagaacataatgacaaaacagccatt

aaaaatatggaacagtttttacgtgagcaatttagtcagttagactggctatccaaccaa
 catattgcgctcgttaggagtaggtggttctgcacgtaattgtgcacgcattcatcaatca
 gcacatgcataccctatcggtagcggttcataactatacgatgacttcgaaagatattaac
 aatgtttatgatctaataccgtaaaagtctcgcgatgaactaacaatttagacgggtta
 5 agtcgcgatcgctcgatattattctgcccgtatctccgtctttaaacacttttcaaa
 aaaattgcagccacacaattcaccttctcaagaaaaggaattcgtgaaggatttattatg
 aaccacatcagcaaacgatatcctgatgagtttaataaaaagtaacgtacgtaagatgca
 ttacgtcatttagcgaacgaatatcatattgaagaaacgagtgctaatacgtcgtgtgaaa
 ttagctcaatccttattgaatcaaattataagtgaaacgatcacttaatttcagaaatg
 10 gaaaaagaattatttattgaaggtgcctacatttattatctaggtagtttcattgattca
 gactcaagttcaccacatacgtattacttaacgcaaatcaatgattaacggcttttca
 cataaagatcgtgtgaaattagctttgttagctagttttaaaaacaaatctttacttaaa
 ttttattgcaaaagaaacacagtggttttagtaataaagaaatagatacaatacaagcttta
 ggagggatatttaaatttgcaaacaccttgaatatctcacatactagttttgtagaggaa
 15 gttaaactaaaagcaaaagaaagatgacaaatacgatttattagttttattacaaaggttca
 cctattgcagaagaataccaagcaaatcgtcagaaaaagcatattgagaaaaattttaa
 ggtaagggtttctattatatttacaacaaatcttaa

Sequence 2860

20 MEERIGLIDIGSNTIRLVIFGYNKKTGLNEILNIKTPARLSQYLTKSNEMNDEGIHVLKE
 TLSSFRKVADKFNVDALYPIATAAIRQSKNREAIKEIKQDIHIEIQIVPEEDEAFYGY
 AITHTTDIENGISVDIGGGSTEVTLEFKDKQLKEAHSPFVSVLKRQFFGDKEHNDKTAI
 KNMEQFLREQFSQLDWLSNQHIALVGVGGSARNVARIHQSAHAYPIGGVHNYTMSKDIN
 NVYDLIRKSSRDELTLNLDGLSRDRVDIILPAISVFKTLFKKIDATQFTFSRKGIREGFIM
 25 NHISKRYPDFNKSNNVRKDALRHLANEYHIEETSANRRVKLAQSLNQLIISERSLNISEM
 EKELFIEGAYIIYLGSFIDSSSPHTYYLIANSMINGFSHKDRVKLALLASFKNKSLK
 FYCKETQWFSNKEIDTIQALGGIIFANTLNISHTSFVEEVKLAKKDDKYDLLVYKGS
 PIAEEYQANRQKKHIEKILKGKVSIIFTKS*

Sequence 2861

Contig_0471_pos_6659_4476

is similar to (with p-value 1.0e-24)

>gp:gp|AF083928|AF083928_2 *Vibrio cholerae* polyphosphate ki
 nase (ppk) and exopolyphosphatase (ppx) genes, complete cds.

NID: g3452464.

35 atgaggtgtatgtatagtatgcaaactcgattgggagaaaaagatattaatttaccgcag
 tattacaacaatagggagtttaagttggctagattttaactacagagtattacaagaatca
 tatgataaaaaataatccgttgcttgaaaaacttaactttatttctatcttcagttcaaat
 ttagatgaattctttatggttcgagtggtgggttaaaagaccaagtcaaaatgggat
 40 gacaacactgaaaaataaagcacagatgacgcctcaagaacaacttgatgctattaaatt
 aaaaaatactgactatgtgaacactcaatatcaacggtataacgaattaattaagaatta
 gccaattacgatattgagatggtaaaacctgaagacttatcagatgcattgatagaaaaa
 ttagaaaaagagttcaagtttaagtgtcttgccgacactcactccggttaggtattgatgcg
 tatcatccatttccaaagttaaataataaaaagtttaaatattttgttgatattcgatacg
 45 gaagatgccattaattcagctatcggttcaaattccttcatttaattccacgctttttaact
 ttaaatgaggggtacaaaaacaatacgttgctcatggtagaagatgtaattacgtatttcac
 aattatctattttacaggatacgaagtactaaataacttttactttccgaataacacgta
 gcagatttaaccattcatgaagatggcgctgaggacttgcttatagaaattgaacgtttc
 ttgaaagaacgtaagagtgggtcggtgtacgtttagaattagattgtcgacttctgaa
 50 aaagagaatgtagaatgggttaacgatcaattagaattgaagataatgatatttattat
 ttagatgggtccacttgatttaacattcttatttggttggttgatcatctatctcataag
 ctcaaatatttaacgtatgagaaatataccctcaaccacctagatcattaggcaacaag
 aatctctatcaattatcattagaaagagatatattctccaccatccgatgaatttatt
 gaaccaatagttgactttattcgacaagcagcagatgacccaataacaatcgctatcaaa
 55 caaaccttgatcgagtgagtaaggattcgccgattattaacagcttaaaagaagccgct
 gaaaacggcaagcaagtaacggtgctcgtagaattaaaagcacgctttgatgaagaaat
 aacgtacattgggcacgtatgttagaagatgctggctgtcacgttatttatggtagaca
 catctaaaaacgcatagtaaaaattgcgctagtcgttaaacgatcaacaataaacttacg
 tcatttgttcatttaggcacaggttaactataacgataaaaactgctaaattatacacagat

atgggtatcatcacgacaaaataagatatcgctgaggatgcaattaacttctttaattac
 ttgagtgggtactcaacaaaaccagaatacaataaattgattgtagcaccatacgaatatt
 cgagacgtcttcattgatcgatcgataaaagaaatacgtagtcatttacaacatggtaac
 5 ggtaaaattatgatgaaaatgaactctttaaccgataaaaacgattatcgaaaagctcttc
 gaagcatcccaagcaggcggttaaaatacaactcatcattcgtggtatatgtgtcttaaa
 ccagycattccaggtatttagcgaaaatatagaggttgtagtattgtgggtcggtt:gtct
 gaacattcacgtatttactacttccataataatggtaggcgcataatttatttatcttca
 gctgacgttatgacacgtaatatgattaaacgtgtcgaaatattattccctgttgaagat
 10 aatcaataggacaacgatttagttaactatatgaatttacaattatctgacaacccaaaa
 ggtcgttaccaagatgcacaaggcggtttatcattatgtcgaaaacaattcatctccttta
 aactctcaatcttacttgatgcaagaagcaattaagtatggagaagaactaaaaaaacaa
 tcggtacaaccttctggacaaccggttcattctagacgtggtggttagttggatacgaaaa
 ttaaaaagtacatttaaaagataa

15 Sequence 2862

MRCMYSMQTRLGEKDINLPQYYNNRELSWLDNFYRVLQESYDKNNPLLEKLNFISSN
 LDEFFMVRVAGLKDQVKMGYDKPENKAQMTPOEQLDAIKIKNTDYVNTQYQRYNELIKEL
 ANYDIEMVKPEDLSDALIEKLEKEFKLSVLPITLPLGIDAYHPPFKLNNKSLNIFVDIDT
 EDAINSIVQIPSLIPRFLTLNEGTSKQYVVMVEDVITYFINYLFTGYEVLNTFTFRITRN
 20 ADLTIHEDGAEDLLIEIERFLKERKSGSAVRLELDCRTSEKENVEWLIDQLEIEDNDIYY
 LDGPLDLTFLFGLVDHLSHKLKYLTYEKYTPQPPRSLGNKNIYQLSLERDIFHHPYESF
 EPIVDFIRQAADDPNTIAIKQTLRVSKDSPINSLKEAAENGKQVTVLVELKARFDEEN
 NVHWARMLEDAGCHVIYGMTHLKTHSKIALVVKRINNKLTSFVHLGTGNYNDKTAKLYTD
 MGIITTNKDIAEDAINFFNYLSGYSTKPEYNKLIVAPYDIRDVFIDRIDKEIRSHLQHGN
 25 GKIMMKMNSLTDKTIIEKLFEASQAGVKIQLIRGICCLKPGIPGISENIEVVSIVGRLL
 EHSRIYFFHNNGEAHIYLSADVMTRNMIKRVEILFPVEDKSIGQRLVNYMNLQLSDNQK
 GRYQDAQGVYHYVENNSSPLNSQSYLMQEAIKYGEELKKQSVQPSGQPVHSRRGGSWIRK
 LKSTFKR*

30 Sequence 2863

Contig_0473_pos_5421_6053
 is similar to (with p-value 8.0e-31)
 >gp:gp|D64024|D64024_2 Sulfolobus sp. DNA for 2-oxoacid:fer
 redoxin oxidoreductase subunit alpha and beta, complete cds.
 35 NID: g1565182.

atggcaacaaagatttaacagttatcgcttctggtggtgatggagacggctatgcaata
 ggaatgggacataactattcatgctcttagacgtaatatgaatatgacgtatattgtcatg
 gacaatcaaatatatggattaactaaaggacaaacatcaccttcctcagctaaaggattt
 gtaactaaatcaacacctaaggaaatatagaaaagaatgtagctccattggaattggca
 40 ctgtcctctggtgcaactttttagcacacaaggattctcaagtgatataaaggcattaact
 aaaatgattgaagatgcgattcatcatgatgggtttttcttttgttaattgtttctcacct
 tgtgttacttacaataaagtgaatacttatgactgggtttaaagaacatttaacaagtatc
 gatgatattgagggctatgacatcacagataaacaacttgctatgaaaactgtgctggat
 catgagtcactggttaaaggatcggtttatcaagatacaacaacaccttcttatgaatcg
 45 caaat:ttcagaactagaacatgaggcggttagctaaaagagatattcatattacagaagaa
 actttcaacgatttaactgcacaatttttataa

Sequence 2864

MANKDLTVIASGGDGDGYAIGMGHTIHALRRNMNMTYIVMDNQIYGLTKGQTSPPSAKGF
 50 VTKSTPKGNIEKNVAPLELALSSGATFVAQGFSSDIKALTKMIEDAIHHDGFSFVNVFSP
 CVTYNKVNTYDWFKEHLTSIDDIIEGYDITDKQLAMKTVLDHESLVKGIVYQDTTTPSYES
 QISELEHEALAKRDIHITEETFNDLTAQFL*

Sequence 2865

55 Contig_0476_pos_11219_12265
 is similar to (with p-value 3.0e-44)
 >sp:sp|P36649|YACK_ECOLI PROBABLE 53.4 KD BLUE-COPPER PROTE
 IN YACQ PRECURSOR. >gp:gp|AE000121|AE000121_8 Escherichia co
 li K-12 MG1655 section 11 of 400 of the complete genome. NID

: g1786306.

atgtataataaagtttttgcaattttaattataattttttccataataattattgcgctct
aatgatacttttcgcagaaagtaagaatgatatgatgaatatgaaagaagataagaaaaat
acaatggatatgacaaatatgaaacatcatgacgaaagaaagaaattaaattcttcacaa
5 ggaaaaaatgaaataatatttcctaaagttgcagagtcaaaaaaagataacaatggttat
aaaaattatacattaaaagctcaggaaggaaagacagagttttacaaaaataatttttct
aatactctaggctacaatggaaatttacttggaccaactttaaaattaaaaaaaggagat
aaagttaaaattaagttaataaataacttagatgaaaatacaacatttcattggcatgga
ttagaagtaaatggaaaagtgatggagggccttctcagggtataaaaccaggaaaagaa
10 aaaactataaaatttgaggttaatcaagattctgctacgttatgggtatcacccccacccc
tctccaaatacagctaaacaagtttataatggcttatcaggattattatatatagaagat
agtaaaaaagaataattatcctagtgtatttgaaaaaatgatttgcctataataatccaa
gataaaaacatttgatatctaaaaaattaaattattcaaaaacgaaagacgaagatggcact
caaggtgatactgttcttgtgaacggaatagtaaaccccaaaactgacaacaaaagaagag
15 aaaatacgtttgagacttttaaatgggttctaatactcgagatttaaatcttaagctaagt
aataatcaaagttttgagtatttgccttcagatggcgggtcaattaaaaaacgctaaaaaa
ttaaaagaaatttaatttagctccttcagaaagaaaagaatagtaaatagattttat:taaa
atgaaaggcgagaaaatcagctcgggtgataatgataaaaactgtaattttaccgattagt
aaacaactggtaagtgtgaagatattag

Sequence 2866

MYNKVFAILIIIFSIIIIASNDTFAESKNDMMNMKEDKKNTMDMTNMKHHDERKKLNSSQ
GKNEIIFPKVAESKKDNGYKNYTLKAQEGKTEFYKNNFSNTLGYNLNLGPTLKLKKG
KVKIKLINLNDENTTFHWHGLEVNGKVDGGSQSVIKPGKEKTIKFEVNQDSATLWYHPHP
25 SPNTAKQVYNGLSGLLYIEDSKKNYPSDYGKNDLPPIIQDKTFVSKKLNYSKTKDEGDT
QGDTVLVNGIVNPKLTTEEKIRLRLLNGSNARDLNLKLSNNQSFYIASDGGQLKNAKK
LKEINLAPSERKEIVIDLSKMKGEKISLVDNDKTVILPISKQLVSVRY*

Sequence 2867

Contig_0476_pos_6429_5413
is similar to (with p-value 6.0e-55)
>gp:gp|X97452|ECPAA_12 E.coli paa cluster for phenylacetic
acid degradation. NID: g2764820.
atgtctcttttgatgacgtcatttttaggtaatacggtaggtaattggggggaatttagct
35 agaaaatcattacttgaagcgggatttagattttaaaatacctgggtataacaattgatcgt
caatgtggctcaggtcttgaagcgttatacaagcctgtaggatgggtacaaagtgtgct
ggaacaatataattgcaggtgggtgttgagagtagaccagtagagcaccttgaaaaacaaa
cgtccgcagtcagtttatgaatctgagtttccacaatttttgaacgggcgccttttgca
agagaaggagaagacccttcaatgattgaagcagccgaaaatgtagcgaagaaatatcat
40 atcagtagaaatgaacaagatgactttgcgtatcgagtcacagttggcatcaaaaaat
atgaataacggtaatatattcccaagaaattttaccggttcaaagtgaaggtgaatatatt
aatcaagatgaaagtattaaacctcaacttactctcagaacacttggcagacttaaacca
cttttaaatgaaggaacagtcacagtaggaaatagttgtaagaaaaatgatgggtgcagta
ttactgattgttatggaagaaaatcgggcacgtcaattaggattcacagaagggttaag
45 tttgtgaatagtgaactgtagggtgttcaaccacagtatattaggagtaggtccagtgcca
gcagtaaatcaatttagctcaagaacgattaaactataaatgatataaatgcagtagaa
ttaaatgaagcatttagctctcaagttattgcgagccaacaacagcttaacattcctttg
aataagttgaattgttggggaggagcaattgctacagggcatccatatggtgcaagtgga
gcagcgttagtcacacgtttattttatatgaaacatcaatttagaactatagcaactatg
50 ggaataggtggagggtataggaaatgcagctttatttgaaagatgggtatggaaattag

Sequence 2868

MSLLDDVILNGTVGNGNLARKSLLEAGLDFKIPGITIDRQCGSGLEAVIQACRMVQSGA
GTIYIAGGVESTSRAPWKIKRPOSVYESEFPQFFERAPFAREGEDPSMIEAENVAKKYH
55 ISRNEQDDFAYRSHQLASKNMNNGNISQEIILPFVKGEYFNQDES IKPQLTLRTLGR LKP
LLI:EGTVTVGN SCKKNDGAVLLIVMEENRARQLGFTEGIKFVNSATVGVQPQYLGVSPVP
AVNQLLAQERLTINDINAVELNEAFSSQVIASQQQLNIPLNKLN CWGGAIATGHP:GASG
AALVTRLFLYMKHQFRTIATMGIGGGIGNAALFERWYGN*

Sequence 2869

Contig_0477_pos_254_727

is similar to (with p-value 8.0e-35)

- 5 >gp:gp|U96107|SCU96107_3 Staphylococcus carnosus N5,N10-met
hylene-tetrahydromethanopterin reductase homolog, SceB precur
sor (sceB) and putative transmembrane protein genes, complet
e cds, and putative Na⁺/H⁺ antiporter NhaC (nhaC) gene, part
ial cds. NID: g2735503.
- 10 atgaaacaaatcaaaacaatctcgacattggtagctggacttggtatagcatttctaggt
cacacaacacatgcagatgcggctgaaaataacaatcaacaacaaagtacatataactat
agtacaactgaagtatcattttctaattcaggaaatttatatacttctggccaatgtact
tggtatgtttatgataaaactggtgaaaaatcggatcaacatgggggaatgcaaatagc
tgggcaactgcagctcaagcagcaggattcactgtaataatacacctgaagaaggtgca
attatgcaatcatctgaaggtgctttcgacatgttgctttcgttgaaagtgtaataat
15 gatggttctattactgtatcagaaatgaactatgatgggtggtccattcgctataagcaca
cgaacaatctctgccagtgaagcaagttcatataattacatccacctgaattaa

Sequence 2870

- 20 MKKIKTISTLVAGLGIAFLGHTTHADAAENNNQQQSTYNYSTTEVSFSNSGNLYTSGQCT
WYVYDKTGGKIGSTWGNANSWATAAQAGFTVNNTPEEGAIMQSSEGAFGHVAFVESVNN
DGSITVSEMNYDGGPFAISTRITISASEASSYNYIHLN*

Sequence 2871

Contig_0478_pos_6998_6654

- 25 is similar to (with p-value 1.0e-60)

- >pir:pir|I67760|I67760 transposase (insertion sequence IS10
) - Escherichia coli >gp:gp|S67119|S67119_2 BST=somatotropin
...BST/beta-Gal fusion protein [Escherichia coli, LBB84 pla
smid pXT107, IS10L/R-1, PlasmidTransposonInsertionMutant, 3
30 genes, 1679 nt]. NID: g455674.
- atgcagattgaagaaaccttccgagacttgaaaagtctgctacggactaggcctacgc
catagccgaacgagcagctcagagcggttttgatatcatgctgctaatacgccctgatgctt
caactaacatggttggttgcggcggttcattgctcagaaacaaggttgggacaagcacttc
caggctaacacagtcagaaatcgaaacgtactctcaacagttcgcttaggcattggaagtt
35 ttgcggcattctggctacacaataacaagggaagacttactcgtggctgcaaccctacta
gctcaaaatttattcacacatggttacgctttggggaaattatga

Sequence 2872

- 40 MQIEETFRDLKSPAYGLGLRHSRTSSSERFDIMLLIALMLQLTCWLAGVHAQKQGWDKHF
QANTVRNRNVLSTVRLGMEVLRHSGYTITREDLLVAATLLAQNLFTHGYALGKL*

Sequence 2873

Contig_0480_pos_5372_6550

is similar to (with p-value 2.0e-24)

- 45 >sp:sp|P23524|YHAD_ECOLI HYPOTHETICAL 42.1 KD PROTEIN IN RN
PB-SOHA INTERGENIC REGION (ORF 3) (F408). >pir:pir|JQ0614|JQ
0614 hypothetical 42K protein - Escherichia coli >gp:gp|D902
12|ECORNPBW_3 E.coli rnpB gene and ORFs. NID: g216630. >gp:g
p|U18997|ECOUW67_54 Escherichia coli K-12 chromosomal region
50 from 67.4 to 76.0 minutes. NID: g606010. >gp:gp|AE000394|AE
000394_2 Escherichia coli K-12 MG1655 section 284 of 400 of
the complete genome. NID: g2367197.
- atgttttaaataatttttggaaaagagaaaaataaggtggttaagacaatgaaagtttta
gtagccatggatgaatttaattgaattattttctagttaccaagctaatagatattgttgaa
55 gaagcggtagcaagtcgaattgaagatgcagatatcggtcaagttccactatttaacggt
cgtcacgaattattagattcagtcctttctttggcaatcaggaaataaatatcggtgtgagt
gcgcatgatgctgacatgaaagaaaccgaagcaatatatggacaaacggatagtggtatg
actattatcgaaggtcacttatttttaaatggcaaaaacatttcaacatcgatcaagtt
tacggtttgggagaggttataaaagcagcattggacaatcatacagaacatcttggtatt

tcttttaggtggaataggaagttttgatggcggtgcaggcatgttgcaagcattgggtgca
 acattttatgatgatgaagcacaattgtcgatatgaggaaaggatgcatatttaataaaa
 tatattagacgtattgatttatcaggtgttcatccacaattaacaaaggtaaacattcaa
 ttaatgtcagattttctcaagtcgattgtatgggaaaaaaagtgaatcatgcaaacatac
 5 gaatcattagattttgtctcaaaatgaagcagccgagatagataatttaatttggtatttt
 agtgaattatttaagaatgaattgaaaatagcaatgggaccaatcgagcgcggtggtgct
 ggaggtggtatagcagctgtattaaatagtctctatcaagctgagattttaacaagccat
 gaattagtgaatcaaatcacacatttagaaaacttaattcaacaggcagatcttattatt
 ttcggagaaggtttgaaagaagaagatcaaattctagagactacaacaatacgtatagca
 10 gaacttaccacaacaatacagcaagccagctattgcaattttgtgctacaaatgataaattt
 gatttgtttgaatcattgaatgttacagcaatgtttaatacatttattgatatgcctgat
 tcatatacagatttttaagatgggtattcaaatcagacattacacagtacaagcactaaaa
 ttattgaaaacgcaataaaacttaccgctttcatcctaa

15 Sequence 2874

MFKIIFGKEKNKVVKTMKVLVAMDEFNGIISYQANRYVEEAVASQIEDADIVQVPLFNG
 RHELLDSVFLWQSGNKYRVSAHDADMKETEAIYQGTDSGMTIIEGHLFLNGKKPIQHRSS
 YGLGEVIKAALDNHTEHLVISLGGIGSFDGGAGMLQALGATFYDDEAQIVDMRKAYLIK
 YIRRIDLSGVHPQLTKVNIQLMSDFSSRLYKKSEIMQTYESLDLSQNEAAEIDNI.IWYF
 20 SELFFKINLKIAMGPIERGGAGGGIAAVLNSLYQAEILTSHELVNQITHLENLIQQDLII
 FGEGLKEEDQILETTTIRIAELTQQYSKPAIAICATNDKFDLFESLNVTFAMFNTFIDMPD
 SYTDFKMGIQIRHYTVQALKLLKTQINLPLSS*

Sequence 2875

25 Contig_0483_pos_5289_6392

>pir:|G64047|G64047 cystathionine gamma-synthase (metB)
 homolog - Haemophilus influenzae (strain Rd KW20) >gp:gp|U32
 694|U32694_2 Haemophilus influenzae Rd section 9 of 163 of t
 he complete genome. NID: g1573035.

30 atgaaggatacagatttagctcaaattgctttaacacaagatcacactggcgcaattgcc
 aatccaatatatcttactgcatatcagcatcctcacctaggatgaatcaacaggctat
 gattatacacgaactaaaaatccaacaagaacagcctttgaagaagcctttgcacaactt
 gaaaaaggcattgcttcatttgcctacttccagtggtatggcgcgattcagttaatatgt
 aatatattcaaacaggatgatgaaattctcgttgcatattgacctatatggtggaacattt
 35 cggttattcgatttttacgaaaaacaatatggtttgaagtttaaatatgtagacttttta
 aattatgaagaagtggaaaaaacatcactccacaaactagagcattatttattgaacca
 atctcaaatccacaaatgattgaaattgatgtagaacatattatccttagcaaaaaa
 catcaactattaacaattatcgacaacactttttaacaccttatctttcgacaccactc
 gaaggagtgcagatatcggttctgcattcagcaacaaaatatattggcggacataacgat
 40 gtgtagctggagttgtaactgttaaggatgctcaattagctgaacaattgaatcaattc
 cataatatgattggagcaactctatcacctcttgatagttatcttttacaagagggtcta
 aagacattacatcttcgcatagagcggttcccaagaaaatgctcaaaaacttgcaaacga
 tgcgccagtcagattcaattgatgaagttttatatagtggaacgaacaggcatgcttagt
 ttaagactgaaccaagcatatagcgctcgctaaatttttagaaaatttagaaatttgata
 45 tttgcagaaagcttaggtggtacagaaacatttatcaccttcccttatacacaaacacac
 gttgatatgccagatgaggaaaaagacaaacgtggaattgatgaatatctcatcagattg
 tccgtaggatagaagactataacgatatagaagctgacataattcaagcattagagaat
 tctaaagtaggagtgatttcatga

50 Sequence 2876

MKDIDLAIQIALTQDHTGAIANPIYLSLAYQHPHLGESTGYDYTRTKNPTRTAFEEAFAQL
 EKGIAFATSSGMAAIQLICNIFKPGDEILVAFDLYGGTFRFLDFYKQYGLKFKYVDFL
 NYEEVEKNITPQTRALFIEPISNPQMIEIDVEPYIILSKKHQLLTIIDNTFLTPLYLSTPL
 EEGADIVLHSATKYIGHNDVLAVVTVKDAQLAEQLNQFHNMGATLSPLDSYLLQRL
 55 KTLHLRIERSQENAKLAQRQRSDSIDVLYSGRTGMLSLRLNQAYSVAKFLENLEICI
 FAESLGGTETFITFPYTQTHVDMPEEDKRGIDEYLIRLSVGIEDYNDIEADIIQALEN
 SKVGVIS*

Sequence 2877

Contig_0484_pos_4823_4407

>sp:sp|069282|MQO_CORGL MALATE:QUINONE OXIDOREDUCTASE (EC 1.1.99.16) (MALATE DEHYDROGENASE (ACCEPTOR)) (MQO).

5 atgataggtactatgattgaaacgcctagagcatgcttgattgcgaatgaccttgcgaaa
cattgtgatttcttcagttttggtactaatgatttaacgcaattgacatttggttctct
agagatgatgcaggaaaattcataaatgtgtatactgaaaataacattttacagcttgac
ccattccaaacttttagatagagaaggtgtaggacgactaattcaattagctgttgaacaa
gctaaaaatacaaatccagagataaaaattggtgtatgtggtgagcttggtggcgatgca
10 aatcaattcgtaaatttaaccaatgggaaattgattatgtttcatgttcaccatttaga
gttccgggtgcaatttttagctacagctcagagtcaagcggaggaaagcgagcgataa

Sequence 2878

MIGTMIETPRACLIANDLAKHCDDFFSFGTNDLTQLTFGFSRDDAGKFINVYTENNILQLD
PFQTLDEGVGRLIQLAVEQAKNTNPEIKIGVCSELGGDAKSIRKFNQWEIDVSCSPFR
15 VPGAILATAQSQAESER*

Sequence 2879

Contig_0484_pos_3408_1879

is similar to (with p-value 6.0e-42)

20 >pir:pir|S53297|S53297 pyruvate,orthophosphate dikinase (EC
2.7.9.1) - Flaveria pringlei >gp:gp|X75516|FPPDK_1 F.pringl
ei mRNA for pyruvate, orthophosphate dikinase. NID: g577775.

atgataactactacacaagaggtgaacattatggctatgtctgacaaaaagacgtcgtg
25 ttaatcggtgctggtgtactaagtactacatttgggtctatgttgaaaacgattgcacct
gattgggacattcatttatatgaacgtctagatcgctcctggtattgaaagttcaaatgaa
cgtaacaatgcaggaaacaggacatgcagctttatgtgaattgaactatactgtacaacaa
cctgatgggttcaattgatattgaaaaagctaaagaaattaatgaacaatttgaaatttct
aaacaattctggggctcatttagttaaatcaggagaaattcaaaatcctaagaatttatt
30 aatccattacctcatattagttttgttcgtggttaaaaaataacggttaaatctttaaagat
cgttatgaagcgatgaagcaattccctatgttcgataatatcgaaatatactgaagattat
gaagaaatgagaaaatggattccattaatgatgaaaggccgtgaagataagggctacatg
gcagcgagtataatagacgaaggaactgacgtaaactacggtgaattaaactcgtaaaatg
gctcaaaatcttaaaaactcaccaaacgttgaagtgaatacaaacatgaagttgttgat
35 tttgaacggtttgtctaatggtaaatggtcagttaaaattaaaaatctaaataatggacaa
gtattcgaacatcaaaactgattatgtgtttatcggtgctggtggcgcaattccacta
ttacaaaaaactggcattccagaaagtaaacatttaggtggattcccaatcagtggtcaa
tttattgcttgtaacaaatccgcaagttattgaacaacacgatgccaaagtttatggtaaa
gaaccacctggtacaccaccaatgacggtacctcacttagatacacggtacattgatggt
40 gaagaacattattatttggaccatttgcacacgtgggacctaattcctaataaacatggt
tctaacttggatttattcaaatcaattaaacatataacattacgactttacttgcttct
gcagttaaaaacttaccattaattaagtattcatttgaccaagtcacatgacaaaagaa
ggttgatgaaccatttacgtacgttctatcctgaagcacgtgatgaagattggcaagtt
tatacagctggtaaacgtgtacaagttattaaagatactgaagagaatggtaaaggattt
45 atccaatttgggtactgaagtggttaattctgaagaccactcagttattgcattactaggc
gaatcaccaggagcatcaacttcagtatcagtagcactagaagttttagagaaaaacttc
cctgaatacgcgaaagactgggaacctaaaatcaagaaaatgattccttcatatggtgaa
tcacttattgatgatgttcaattaatgagaaaaatacgtaaacaaacatctaaagattct
gaattaggattctataataaagcaaaaataa

Sequence 2880

MITTTQEVNIMAMSOKKDVVLIGAGVLSTTFGSMLKTIAPDWDIHLYERLDRPGIESSNE
RNNAGTGHAALCELNYTVQQPDGSIIDIEKAKEINEQFEISKQFWGHLVKSGETIQNPKEFI
NPLPHISFVRGKNNVFLKDRYEAMKQFPMFDNIEYTEDIEEMRKWIPLMMKGREDKGYM
55 AASKIDEGTDVNYGELTRKMAQNLKNSPNVEVQYKHEVVDFERLSNGKWSVKIKNLNNGQ
VFEHQTDYVFIGAGGAIPLLQKTGIPESKHLGGFPISQFIACTNPQVIEQHDAKVYVK
EPPGTPPMTVPHLDTRYIDGERTLLFGPFANVGPFLKHGNSLDLFSIKPYNITLLAS
AVKNLPLIKYSFDQVIMTKEGCMNHLRTFYPEARDEDWQVYTAGKRVQVIKDEENGKGF
IQFGTEVVNSDHSVIALLGESPGASTSVSVALEVLEKNFPEYAKDWEPIKKMIPSYGE

SLIDDVQLMRKIRKQTSKDLELGFYNKAK*

Sequence 2881

Contig_0486_pos_1073_1840

5 is similar to (with p-value 9.0e-58)

>gp:gp|U38892|SSU38892_2 Synechocystis sp. ruvB gene, complete cds, and secA gene, partial cds. NID: g1256587.

atgttttaaaataggaatttagaattacaatctcgtttacttttaggtactggaaaattt
gaaaatgaagaggttcagtcaaaagcaattgagcatctgaaacaaatgtacttacattt
10 gcagtaagacgtatgaatttatatgatcgtaacctacctaaccacttgcaaacgttaat
ttaaagatttttactcattttccaaatactgcaggtgccaaaacagctcaagaagctatc
agaattgctgaaattgctaatacgcaggtgtatgtgacatgattaaagtcgaagtcatt
ggtgatgacgaaacattattacctgatccattcgaaacatacgaggcatgcaagatttg
ttagaaaaaggttacactgtttgtccttacatctctaacgatttagtttagctcaacgt
15 ttagaagaattgggtgtacacgcagttatgccacttgcatcccctattggtacaggaaga
ggtattaataaccattaaatttaagttatattatcgaaaatgctagtgtacctgtaac
gtagatgctggtattggttcccctaaagatgcgtgtcatgccatggagcttggcgagat
ggtattttactcaacacagcatttcagcggcaaaagatcctgtgaaaatggctgaagca
atgaaattaggatataaatgctggcagactttcatatgaagctggacgattcctgttaag
20 tatactgcacaagcatctagtccatcagaaggtttagggttctttaa

Sequence 2882

MFKIGNLELQSRLLLTGKGFENEVQSKAIEASETNVLTFAVRRMNLVDRNLPNPLANVN
LKDFITFPNTAGAKTAQEAIRIAEIANHAGVCDMIKVEVIGDDETLPPDFETYEACKVL
25 LEKGYTVCPYISNDLVLAQRLEELGVHVMPLASPIGTGRGINNPLNLSYIIENASVPVI
VDAGIGSPKDACHAMELGADGILLNTAISAAKDPVKMAEAMKLGINAGRLSYEAGRIPVK
YTAQASSPSEGLGFL*

Sequence 2883

30 Contig_0487_pos_1335_2006

is similar to (with p-value 5.0e-34)

>sp:sp|P39762|AMPS_BACSU AMINOPEPTIDASE AMPS (EC 3.4.11.-).

>gp:gp|AF012285|AF012285_21 Bacillus subtilis mobA-nprE gene region. NID: g3282109. >gp:gp|Z99111|BSUB0008_117 Bacillus
35 subtilis complete genome (section 8 of 21): from 1394791 to 1603020. NID: g2633699.

atgacgaattatcataataagttaaaacaatatgcagaattattagtaagagtgggaatg
aatgtacaaccacagcaacctgtttttatcgttcacatctgttgaaagcgttagaattaact
catttaactcgtcgaggaagcatataaagcaggggcagaagatgttcgagtgagctacaca
40 gacccgaaattaaaagattaaaatttgaaaacgaatcagttgaacactttgaaaaacaa
gaactcaacaatatgatattgaagagcgtctggattatgttaactcgtggcgagcgaac
ttggcgctcattgctgaagatccagagctattaaatggaatagatgcgcaaaagttaaaa
gcgatcaaaactgtatactcaaaaggatttaaacatatatggaagcaagtcaaaaaaac
caatttccatgggtagtggctgcgttccctactagggttgggcacgtcgtgtctatcca
45 gagttggatgttgaaatcagcatatattaaattcattgatgaagtatttgatattgttcgt
gtagatggacaaaatccaattgaaaattgggaaaaacacattaaagatttaagtgttcac
gctaaacgattacaagagaaaaactatcaagctttacattacatatcagaaaattcatac
atttggtttga

50 Sequence 2884

MTNYHNKQYAEELLVRVGMNVQPPQVFIRSSVEALELTHLIVEEAYKAGAEDVRVSYT
DPKLKRLKFENESVEHFEKQELKQYDIEERLDYVNRGAANLALIAEDPELLNGIDAQKLK
AYQTVYSKGFKPYMEASQKNQFPWVVAAPTRDWARRVPELDESAYIKFIDEVFDIVR
VDGQNPIENWEKHIDLSVHAKRLQEKNYQALHYISENSYIWF*

55

Sequence 2885

Contig_0487_pos_3511_3086

is similar to (with p-value 6.0e-23)

>sp:sp|P14597|DUT_ORFN2 DEOXYURIDINE 5'-TRIPHOSPHATE NUCLEO

TIDOHYDROLASE (EC 3.6.1.23) (DUTPASE) (DUTP PYROPHOSPHATASE)
 . >gp:gp|M30023|ORFPRTS_1 Orf virus homologue of retroviral
 pseudoprotease gene, complete cds. NID: g332561.

5 atgacaaatacattagaaattaaattattatcagaaaacgcgactatgccgaagagagca
 aattctacagatagtgattggactttatcgtatcagaaacgattaacattcctgcacac
 gcaactaaagtagttaaaacagatatagcgattaatctgccttatgggtatgaggcgcaa
 gtaagacctagatctggtaaatcacttaaaactaaattgcgtgtagcactaggaacaata
 gaccaaatacaccacaaagaataggtatcatcacagataacataggtaatgaagatatc
 acagtagaaaaaggagaaagattagcgcaattagttgtagcgccagttgtatatcctaca
 10 cccaaacaggttgattggtttgaaaatgaaagcgacagaggtgcatatggaagcacagga
 gaataa

Sequence 2886

15 MTHPLELKLSENATMPKRANSTDSGLDLYVSETINIPAHATKVVKTDIAINLPYGYEAQ
 VRPRSGKSLKTLRVALGTIDQTYHKEIGIITDNIGNEDITVEKGERLAQLVVPVYPT
 PKQVDWFFENESDRGAYGSTGE*

Sequence 2887

Contig_0487_pos_988_515

20 is similar to (with p-value 2.0e-17)
 >sp:sp|P41893|PPAL_SCHPO LOW MOLECULAR WEIGHT PHOSPHOTYROSI
 NE PROTEIN PHOSPHATASE (EC 3.1.3.48) (LOW MOLECULAR WEIGHT C
 YTOSOLIC ACID PHOSPHATASE) (EC 3.1.3.2) (PTPASE) (SMALL TYRO
 SINE PHOSPHATASE). >pir:pir|A55446|A55446 protein-tyrosine-p
 25 hosphatase (EC 3.1.3.48), low molecular weight - fission yea
 st (Schizosaccharomyces pombe) >gp:gp|L33929|YSPLMPTP_1 Schi
 zosaccharomyces pombe low Mr protein tyrosine phosphatase mR
 NA, complete cds. NID: g602991.
 gtgataactaatgatacatgtagcatttgtatgtctcggaatatatgtcgttctccaatg
 30 gctgaggctatcatgagacaaagactacaagaaagaggtatttcagatatataaagttcat
 tctagagggaacaggagcgttggaatttaggcgaacctccacataacggaacacaaaaaatt
 ctacagaagtaccatattccttatgatggtatggtgagtgaacttttcgaacctgatgat
 gattttgactatattattgctatggaccaaagtaacgtagacaatatcaaacaaatcaat
 ccaaaatttacaaggacaattgttcaaattgctagaatttagtaacatggaagagagtgat
 35 gtaccagatccatactacacaaataattttgaagggtgttttcgagatggtgcaatcatct
 tgtgataatttaatagactacatcgtaaaagatgcaaatttgaaagagaggttaa

Sequence 2888

40 VILMIHVAFVCLGNICRSPMAEAIMRQRLQERGISDIKVHSRGTGRWNLGEPHNGTQKI
 LQKYHIPYDGMVSELFEPDDDFDYIIAMDQSNVDNIKQINPNLQGQLFKLLEFSNMEESD
 VPDPPYTNFEGVFEMVQSSCDNLIDYIVKDANLKER*

Sequence 2889

Contig_0488_pos_5255_6256

45 >sp:sp|P44770|OTC_HAEIN ORNITHINE CARBAMOYLTRANSFERASE (EC
 2.1.3.3) (OTCASE). >pir:pir|H64079|H64079 ornithine carbamoyl
 ltransferase (arcB) homolog - Haemophilus influenzae (strain
 Rd KW20) >gp:gp|U32741|U32741_4 Haemophilus influenzae Rd s
 ection 56 of 163 of the complete genome. NID: g1573582.
 50 atgaaaaatttacgtaacagaagctttttaactttattagacttttcacgacaagaggt
 gaatttttattaacactctccgaagatttgaagcgtgccaaatataatcggcactgaaaag
 cctatgctaaaaaataaaaaatatcgcgcttcttttgaaaaagattccactagaacacgt
 tgcgcattcgaaagttgccgcacatgatcaaggtgcacacgtcacttatcttgaccaca
 gggtctcaaatgggtaaaaaagaaactgctaaagatacagcacgtgtacttggtggtatg
 55 tatgatggtattgagtaccgaggtttctctcaacgtactgtagaaacattagcgcaatat
 tcagggtgtccggtatggaatggattaaccgatgaagatcacctacacaagtgttgct
 gactttttaactgctaaagaagtattgaaaaaagagtatgctgatatacaactttacttat
 gttggcgatggacgtaacaatgttgtaacgcattaatgcaagggtgctgccattatgggt
 atgaatttccatcttgtttgtcctaagaactcaatccgacagaagaattattaatcgt

5 tgcgaacgtattgcgacggaaaatggcggtaacattttaataacagatgatattgataaa
 ggcgtgaaagattctgatgtattttatacagatgtttgggtatcaatggcgaaacctgat
 gaagatggcaagaacgccttaaaacttttaaaaccatatcaagttaaccaagcattatta
 gaaaaaacaggcaatccaaatgttatttttgaacattgtttaccttctttccacaatgca
 gaaactaaaattgggtcaacaaatttatgaaaaatatggcattagtgaatggaagtcact
 gatgatgtcttcgaaagcaaaagcttctgtagtattccaagaagctgagaatagaatgcat
 acaattaaagcgtcatggtagcaacttttaggagaattctaa

Sequence 2890

10 MKNLRNRSFLTLDFSRQVEVEFLTLSEDLKRAKYIGTEKPLKNKNIALLFKEDSTRTR
 CAFEVAHNDQGAHVITYLGP TGSQMGGKETAKDTARVLGGMYDGI EYRGFSQRTVETLAQY
 SGVPVWNGLTDEDHPTQVLADFLTAK EVLKKEYADINFTYVGDGRNNVANALMQGAAIMG
 MNFHLVCPKELNPTEELLNRCER IATENGGN ILITDDIDKG VKDS DVIYTDVWVSMGEPD
 EVWQERLKLKPYQVNQAL EKTGNPNVIFEHCLPSFHNAETKIGQQIYEKYGISEMEVT
 15 DDVFESKASVVFQEAENRMHTIKAVMVATLGEF*

Sequence 2891

Contig_0488_pos_6270_7211

is similar to (with p-value 3.0e-76)

20 >sp:sp|Q46807|ARCL_ECOLI CARBAMATE KINASE-LIKE PROTEIN 1. >
 gp:gp|U28375|ECU28375_24 Escherichia coli K-12 genome; appro
 ximately 64 to 65 minutes. NID: g887800. >gp:gp|AE000370|AE0
 00370_9 Escherichia coli K-12 MG1655 section 260 of 400 of t
 he complete genome. NID: g2367170.
 25 gtgagtgaatggctaaaattgtagtagctttaggtgaaacgctttaggaaaatcacc
 caagaacaacttgaattagtaaaaaatacagctaaatccctagtaggattaattactaaa
 ggtcacgaaattgtgatttagtcacggtaatggaccacaagtaggaagtattaaccttgg
 ctgaattatgcagctgaacacgatcaaggtcctgcttttccatttgcgtgaatgtggcgct
 atgagtcaggcctacatcggtatcaacttcaagaaagttacaaaatgaacttcattca
 30 atgggcatagataagcaagttgtcacactagttacccaagtagaagttgatgaaggcgat
 ccagcttttaatatgtccaagtaaacccatcggtctgttctacactaaagaagaagcaaat
 cgtattcaacaggaaaaaggttatcaattttagaagatgctggctcgaggttacgcgcgc
 gttgtaccatcaccacaaccaaatatctattatcgaactggaagattaaaaactctagta
 gaaaatgacacactcgtcatcgctgcaggtggaggtggtataccagtcattcgcgaacag
 35 catgatagctttaaaggtatagatgccgtcatcgataaagacaaaacaagtcattatta
 ggtgctgatattcactgtgatcaactcattattttaacagcgattgattatgtttatatac
 aactatcatactgaccaacaacaagcacttaaaacaacaatatagatacgcttaaaaca
 tatattgaagaagaacaatttgccaaaggcagcatgctacctaaaatcgaatctgccatc
 tcctttattgaaaaataatcctaacggtagcgtgctcatcacatcattaaatcaattagat
 40 gcagcactagaaggtaaaattggcacactcattacaaagtaa

Sequence 2892

VSEMAKIVVALGGNALGKSPQEQL ELVLKNTAKSLVGLITKGHEIVISHGNGPQVGSINLG
 LNYAAEHDQGPAPFPAECGAMSQAYIGYQLQESLQNELHSMGIDKQVVTLVQVEVDEGD
 45 PAFNSPSKPIGLFYTKEEANRIQQEKGYQFVEDAGRGYRRVVPSPQPI SII ELESIKTLV
 ENDTLVIAAGGGGIPVIREQHDSEFKGIDAVIDKD KTSALLGADIHCDQLIILTAIDYVYI
 NYHTDQQQALKTTNIDTLKTYIEEEQFAKGSMLPKIESAISFIENNPNGSVLITSLNQLD
 AALEGKIGTLITK*

Sequence 2893

Contig_0489_pos_5066_5410

is similar to (with p-value 5.0e-39)

50 >sp:sp|P37941|ODBB_BACSU 2-OXOISOVALERATE DEHYDROGENASE BET
 A SUBUNIT (EC 1.2.4.4) (BRANCHED- CHAIN ALPHA-KETO ACID DEHY
 55 DROGENASE COMPONENT BETA CHAIN (E1)) (BCKDH E1-BETA). >pir:p
 ir|S32487|S32487 3-methyl-2-oxobutanoate dehydrogenase (lipo
 amide) (EC 1.2.4.4) E1 beta chain - Bacillus subtilis >gp:gp
 |M97391|BACBRANCH_2 Bacillus subtilis branched chain alpha-k
 eto acid dehydrogenase E1-alpha, branched chain alpha-keto a

cid dehydrogenase E1-beta, and branched chain alpha-keto acid dehydrogenase E2, complete cds. NID: g142610. >gp:gp|D84432|BACJH642_247 Bacillus subtilis DNA, 283 Kb region containing skin element. NID: g2627063. >gp:gp|Z99116|BSUB0013_115 Bacillus subtilis complete genome (section 13 of 21): from 2395261 to 2613730. NID: g2634723.

gtgaattactgtttacaagctgcagatattttggcaaatgacggcatcgatgttgaagtagtcgacttaagaacagtttatccactagataaagcaactatcattgaacgctctcaacgtagtggttaaagttcttctgttactgaagataatctagagggaagcattatgtctgaagtagtctgcaattatagctgaaaactgtctgttcgatttagatgcgccaatcatgcgattagctgcaccggatgtcccatctatgccattttcaccaacattagaaaatgaaattatgatgaaccagaaaaagatacaggacaaaatgcgtgaactcgcacatttttaa

Sequence 2894

15 VNYCLQAADILANDGIDVEVVDLRTVYPLDKATIIERSQRTGKVLVLTEDNLEGSIMSEV
SAIIAENCLFDLDAPIMRLAAPDVPSMPFSPPTLENEIMMNPEKIQDKMRELAQF*

Sequence 2895

Contig_0493_pos_2490_1000

20 >sp:sp|069282|MQO_CORGL MALATE:QUINONE OXIDOREDUCTASE (EC 1.1.99.16) (MALATE DEHYDROGENASE (ACCEPTOR)) (MQO).
atgcacatgagtgaaacaaatcataaaaacatcggtgtttaggtgcaggaattattggtagcgtcagtagcgacaatgctttcaaaagtaagtcctaactggcatatcgatatgtttgaagactagaaggcgctggtattgaaagttcaaatgaaaataataatgctgggacaggtcatgaggcattatgtgaattaaactatacagttgaacaagatgatggttcaattgatgcattaaagcgcaagaaattaatgaacaattcgaattatctagacaattctggggaatttagttaaaaatggtgatattttctaattcctgaagaatttattcaaccattacctcatatcagtttcgttatgggaccaacaaacggttaactttttaagaaaacggttatgaaacactaagaactcttccaatgttcgatacaatcgaatatacagaagacatggaaacaatgagaaaatggatgcca

30 ttaatgatggaaaatcgtgaaccaggtcatcaaatggcagcaagtaaaattgatgaaggtagacagatgtgaactatggtgcgttaacacgtaagttagcacattacttagaacaaaaatctaatgtttcattaaaatacaatcatgatgtttagatttaacacaaagagaagatggcaaatgggaagttgtcgltgaaaatagagaaaactaaagaaaagtaactaaaatagcagataaa

35 gtgtttattggtgctggcggtcactctattccgttattacaaaaatctggcggttaacaaagagaacacctaggtggtttcccaatcagtggtcaattcttaagatgtacaaacccagatattattaacaacatgcggctaaagtttacagtaaaagagcctcaaggtaagccaccaatgactgtaccacaccttgatacacggttatatcaatggtaaacaaacattattatttggtccatagcgaatatcggccctaaattcttgaaattcggttcaaatctagacttattcgaatcatcaaaacatatataacattactacaatgttggcttcagcagttaaaaatgtacctttaatt

40 aatatattcaattgatcaaatgatcaaaactaaagaaggttgatgaactatttaagaacatttatctctgatgctaaagatgaagattgggaactttacactgctggtaaacgtgttcaagttattaaagatagtgaaacacacgggaaggtttcgtagtagtttggtactgaagttgtcaattcagacgacaattctatgattgcattattaggtgaatctccaggggcttcaacatcatatcagttgtattagaagtttttagagaaaaacttcgctgatgacaaagaagcatgggaa

45 cctgttggttaaagaatggtaccaacatacggctcggttcattaattaatgacgaaaaattaatgagagaaacacgctcgcgaaacttctaaaaacttacatttaaatagataa

Sequence 2896

50 MHMSEANHKNIVVVGAGIIGTSVATMLSKVSPNWHIDMFERLEGAGIESSNENNNAGTGH
AALCELNYTVEQDDGSIDASKAQEINEQFELSRQFWGNLVKNGDISNPEEFIQPLPHISF
VMGPTNVNFLRKRYETLRTLPMFDTIEYTEDMETMRKWMPLMMENREPGHQMAASKIDEG
TDVNYGALTRKLAHYLEQKSNVSLKYNHVDVLTQREDGKWEVVVENRETKEKVTKIADK
VFIGAGGHSIPLLOKSGVKQREHLGGFPISGQFLRCTNPDIKQHAQVYSKEPQGGKPPM
TVPHLDTRYINGKQTLLEGPYANIGPKFLKFGSNLDFESIKPYNITTMLASAVKNVPLI

55 KYSIDQMIKTKEGCMNYLRTFIPDAKDEWELYTAGKRVQVIKDSEQHGKGFVVFGTEVV
NSDDSMALLGESPGASTSLSVVLEVLEKNFADDKEAWEPVVKEMVPTYGRSLINDEKL
MRETRRETsknLHLNR*

Sequence 2897

Contig_0494_pos_12199_12525

>gp:gp|AB009866|AB009866_15 Bacteriophage phi PVL proviral DNA, complete sequence. NID: g3341907.

5 atgaattcagcagtagtagagtcagaactgaattctacaccatcttctttggtattacct
tctaattcaggaatgtgaataatccttttaggtaatccacatactcgcgtaaccatct
tccatagttttcgcaacataacagctacatatgggtggtgtatcggtaccaactgacacg
atgccgtcctctgatttttctaataccgaacaatgcaactctgtcctctaattggtagtgtg
tggaaccagcttctacttcaattgttccgttagcaactgccatttctgcaacttggta
10 tcaccaiatgccttctcgatgtcttga

Sequence 2898

MNSAVVESELNSTPSSSLVLPNSNGVNNPLGNPTYSREPSSIVFANITATYGGVSLPTDT
MPSSDFSNNPNATLSSNGSLWKPASTSIVPLATAISATWLSPLYAFSMS*

15 Sequence 2899

Contig_0494_pos_16099_14756

is similar to (with p-value 3.0e-17)

>gp:gp|AB009866|AB009866_13 Bacteriophage phi PVL proviral DNA, complete sequence. NID: g3341907.

20 atggctaatttagatgagcgcaaaaaagaaatcgctaattctgatttctaagcgcaagaa
gcagtcgaaaaaggcgacctcgaaactgctcgtaatttaaaagctgatattgatgctcaa
aagaaagagtacgaagaactcgaaacagcttcaaaagaaattgaagcgtcagcacctaaa
caagatgaaccacctaagatgaaggtgcagaagttgaagataacaaagatggtaattct
ggagaagaatcagagaacaaaccttctgatgatgaaccagaaggaacttcagatgaagaa
25 aaacctgatgatgcacaaaaccagatgacaaacctgaagaacaccagaaacacctact
attgaaaaagtgagaagaaaccaacagaagaattaaaaaaaagaaaaagacaaaaaagaa
ggagcgaaacggttctatggctaaattaaacaaaatccagagacaaacgaagaaatcta
gcacacgacagtagatgaaatcaaaaggggctaaacgtgacaatgttaaatctgtgac
gttgccgtaactatcccagaggatattaaatatattcctgaaaaagaagttaagacagtc
30 caagacttatcagaattgggtacaaaaaacttcagtatcaactgcaagtggaataccg
atcttaaaacgtgctaacgctaaattcaacactggtgctgaattagagaaaaacctgag
ttagctcgctccggaattcgaaacaatcaattgggaagtagacacttatcggtgatctatt
ccgatttcacaagaagcatttagatgattcagttgctaacttaactgctattgtttctgaa
aatattaacgaacaaaaaatcaacactttaaatgaacgtattgggtgaagttttaaaagca
35 ttcaatcctactagtgtttctaattgttgacgacttaaaagaaattatcaacgttaaat
gatcctgggttatgaccgccaattatctgtactcaaagtttctatcaaaaactagataca
ttaaaagatggtaacggtcggtattttactacaagacagtatcatcaacactgcaggtaac
actgtgttaggtatgaatgtaacagttgtgcgtgatgacttgttaggtaaaaatggagat
gcattagcatttattgggtgatgtaaaacgcggtgtgttatttgacagccgtacagacggt
40 tctgttcaatggattgaaaaatgaaatctacggtaataacttaattgggtgctttccgtttc
gatgtgaaacaggtgatataaatgctggtttcttcgtaacatttgaagatgcaacagaa
cctagtggggatctaggagcataa

Sequence 2900

45 MANLDERKKEIANLISKAEAVEKGDLETARNLKADIDAQKKEYEELEQLSKEIEASAPK
QDEPPKDEGAIVEDNKGNSGEESENKPSDDEPEGTSDEEKPDDAPKPDCKPEETPETPT
IEKVEEPTTEELKKEKDKKEGAKRSMKLNQNPETNEEILAFEQYMKSKGAKRDNVKSDD
VGVTIPEDIKIYIPEKEVKTVQDLSELVQKTSVSTASGKYPILKRANAKFNTVAELEKNPE
LARPEFETINWEVDTYRGSIPISQEALDDSVANLTAIVSENINEQKINTLNERIGEV LKA
50 FNPTSVSNVDDLKEIINVKLDPGYDRQIICTQSFYQKLDTLKDGNGRYLLQDSIINTAGN
TVLGMNVTVVRDDLKNGDALAFIGDVKRGVLFADRTDVSQWIEENIYGKYLMAFRF
DVKQADKNAGFFVTFEDATEPSGDLGA*

Sequence 2901

55 Contig_0494_pos_14069_13677

is similar to (with p-value 5.0e-25)

>gp:gp|AB009866|AB009866_13 Bacteriophage phi PVL proviral DNA, complete sequence. NID: g3341907.

atgaaaggcgataaaagaaataattgcctatttagaaacgaaatacggtaaatctgctatg

5 aaacgcataactgactttgcactaactaaagctggcaataaagttgtaagtattatcaaa
 ggtaatatgaaaagttttgaagatactggagaatcagtagaagaaactacactttcaaa
 ccgatgacgataaaaaggagtaaggaccgttaaaattcattggcgaggtcctaacaacgt
 tatcgtattatccacctaataatgaatatggctactttgatcgttctggaaagtgggttaat
 acagctggtaaaggtgttattgaaaacgctatgcgtgaaggcagagaaacgtatttcaga
 acagtaaaagaagaatgagaaagaggggtgtaa

Sequence 2902

10 MKGDKEIIAYLETKYKGSAMKRITDFALTKAGNKVVSIIKGNMKSFEDTGESVEETLSK
 PMTIKGVRTVKIHWRGPKQRYRIIHLNEYGHFDRSGKWVNTAGKGVNIENAMREGRETYFR
 TVKEEMRKRV*

Sequence 2903

Contig_0494_pos_12659_12030

15 No hits found

atggcagagaaaaactatcgttcatttacagggttaacagaattttattataaagttcat
 ggagaaggtggagttcaaaaagttgctgatccagaacgcattaaatatttacaagaaatt
 tcagtatctaaagatcaagacatcgagaaggcatatggtgataaccaagttgcagaaatg
 gcagttgctaaccggaacaattgaagtagaagctggtttccacaaactaccattagaggac
 20 agagltgcattgttcggattagaaaaatcagaaggacggcatcgtgtcagttggtaacgat
 acaccaccatattgtagctgttatgtttgcgaaaactatggaagatggttcacgcgagtat
 gtgggattacctaaaggattattcacattccctgaattagaaggtaataccaaagaagat
 ggtgtagaattcagttctgactctactactgctgaattcatgcaagctaaagttaaaggc
 ttcgaagaagaaaaagcaatgttattaggtcacgatgctaaaggcacaactgttatgaaa
 25 gacgctatctgggaagctatcttcggtgaatctgcaccaagcagtgatccaaaagaatct
 agtggaaacagaatcagaactaggcgcataa

Sequence 2904

30 MAEKNYRSFTGLTEFYKVVHGEQVQKVADPERIKYLQEI SVSKDQDIEKAYGDNQVAEM
 AVANGTIEVEAGFHKLPLEDRVALFGLEKSEGDIVSVGNTPPYVAVMFAKTMEDGSREY
 VGLPKGLFTFPELEGNTKEDGVEFSSDSTTAEFMQAKVKGFEEKAMLLGHDAKGTVMK
 DAIWEAIFGESAPSSDPKRESSGTESELGA*

Sequence 2905

35 Contig_0494_pos_11140_6422

is similar to (with p-value 1.0e-33)

>gp:gp|AB009866|AB009866_7 Bacteriophage phi PVL proviral D
 NA, complete sequence. NID: g3341907.

40 atgaaagcctatggcggttcaacgtagtatttcggaataaaacgtagctttaaagyttaa
 aacgctgacttaaaactatctaaacaactttaagtattccgaaaaaagtttaattca
 tataagtttaagaactagagaattatcgcaagcagtcataagaatctaaagctaactgtgca
 gcggtgaaagcaaaaataccaagaagcagcaagagcatctggtgtgaatagcaaaaagcc
 gctcaattaaaggcaggaatatagtcgacaagctgacaatctcaactatttacaaaacgaa
 ctcgaccaaacacgtggcaatacagagaaatgattgcagtaagtaaatcatctgtcgggt
 45 agacttgggcaagcattttctgaaataggacctaagataagatccataggagattcaatg
 aagtcagtcgggcgtaacatgagtttacacggttactgcaccaattgcagcaggttttggc
 gctgcggtgaagaaaagttagacttcgacgataccatgcgtaaagtaaaagccacatct
 ggtgctactggagatgaatttaaccagcttagaacaagaagcacttcaaatgggcccagat
 actaaattcacggcctctgaatctgctgaagcaatgaactacatggcgcttgctggttgg
 50 gacaccaagatatgctaaaaggtgttgggtgtaatggatttagctgctgcatctggt
 gaagatttagcaagcgtatctgatattgtaactgataacctaactgcatttggtatgaaa
 gctaaagatagtagccactttgctgatgttttggtcaaacgagttcaaaaagctaatact
 gatgtacgtggttttaggtgatgcgttttaatatgctgctccagttgctggtgcgttaggt
 tacacggtgaagatacatcaatagctatttggttgatgtctaatgctgggataaaaaggt
 55 gaaaaagccggcagcagcattaagaacaatgtttaccaacctaactaaaccaacaaaagca
 atgaaagacgaaatggataaactaggaatatctattactgatagcaacgggtgaaatgta
 cctatgagagatgttttagatcaacttagaggtaaaatggcggtctatctaaagaccaa
 caagcagccgcagctagtagcaatatttggtaaaagaggccatgagtggtgcatttagcagtt
 atcaatgcctacagacgaagattataaaaagctaactaaatccatagacggctctaaggt

gcttcaaaaagaatggctaagaatggaaggcggtattggtggcgcaatgcgtaaaatg
 aaatcggcgaattgaaagtttagcgatttcattaggtgatgcattagcccaatgttatat
 aaagctgctaaatggatcacatcattagcgaataagtttttctaatttacctactggcgtt
 5 caaaaaacgattgcagttgtaggattactcgccgcagctattggtccactactaatggct
 tttggcggttatggcatcaacaattggcactgctataacagttattaggctctttaatgacg
 agtatgagaacactatcatttttatctaaaaccagtgacgacgagcgtggtatttggaat
 ggcgttactgccactgctcgtggtatcgcaaatggttatagatatgcggtggctgcatta
 accacttctcagacaatacaggctatgaaaactaaaattgctgcagctgcaacaacagct
 10 tggactacagttactaaagtgcaacttttagcaactaaaggcttgggattagcaataaga
 tttatgactggaccagtcggtatagttattacagccatcggattattagtagcagggtct
 attcattttatggaaaacaaatagctcgtttagaaatgcagttatcggcatttggaattca
 ataaaaatgctgcaatagctatatttggttttatcaaaccttatattattaatatgttg
 aacgcaattaaaaactctacaattgccatttggaacgagattaaaaagaagtgcgttaata
 atatggaacgctattaaatttgctgttcaacatcctattcaagcattaaaaaatgtctta
 15 tcagctttatggaatggcatgaaaaatgctgctattaaaatctggaccgccttaagaac
 ggtgttatagcaattattaaagcatatgttgcgcaagtaaaatttaatatcaacctatt
 aagcgcatagtagttacgatatttaatgctattaagagtttctctattaaagtttgact
 gcattaaaaaatgggtgtgtaggaatcggttcgagctttgcgcaagggtgttctatctgta
 tttaacgcattaaaaaaaggtattgttgtaattttaatgctgtaaagaatgccgcagtt
 20 aaaatctggacggctataaaaaaatcagtagtgaataaagcaaaagcattatggctgga
 gttaaaaatacatggaatgcactcaaaaaaggtacaattggcataatttaagcagttggc
 agtttcatgagttctaaatggaacagttataaaaaaggtactgttaataaagcgaaagct
 ctatggtcaggcgtcaaagtgcttggggatcacttaaaaaaggtactcataacaccatg
 aatgctgtaggtggcttcatgagcaagaaatggaatggaattaaaagtactactgtatct
 25 atagtaaatggcatgaaatcgaaagttagggcaccatgaataaaatgagagacggatc
 aaaacagttaccggtaaaattgggaatcttttgggtggaatggttaaggcggttaaaaaa
 ggccttaatggattaatcaaaaggtgttaactgggtcgagataaaatagggtatggataaa
 atacctaaagattaaattatctacaggtactcaatctacgcatacacaaagttatattacc
 aacggtaaaatcaataaaaggtacaatggctactgtaggagatagaggtcgaggtaatgga
 30 cctgggtggctttagaaatgaaatgatccgttatcctaacggaagaatggcactaacacct
 aacaagatacaacaacattcttacctaaaggctcaagtgtatatagtgccgcgcgaaca
 catgctattttatctaattcaggatagacactaagaagaaaaaactacctaaatttagt
 aaaggtactaaaaagaaagacgggtatattagatgttattagctctgggtgtaaaaaatgat
 gtttaataaagttaaagacattgggtggttaaagcaagagacataggtgggtactacgtttgac
 35 aaagcaaaagacataggtacaaaagcacttgataaagctaaagatgtgtctagcactgtt
 atcaagggtatggagatgttttgattatgtaggtcacctatgaaattggtaaaataaa
 gtctttgagaaagtgggttttaacctagactttatgaaaaatgcaccattaccatttgat
 ttaatgacagctatgattaaagaaacttaaaaaatgggtattaaagacttctttaatgaaagt
 tttagcttgcaggcgggtggagatgggttcttcgttcactaaattcccaattactacgggg
 40 tattatcctaattgggtgggtgctcctgggttatagttttaatgggtgggtgctcacttgggtatt
 gactatggcgctccatattggtacaactatcaatgctaccaatgatggaaatgtaaaagct
 atccacaacttaggtggaggacttgttgcacgacttttaacaggtcagttcacattgttc
 tttatgcacttatctaaaatattaaaacaaggtaaaatcaaagctggagaaccaatggct
 aaaacaggttaattcaggtcaatggactactggtccacacgtacacttccaagttgaaaga
 45 ggtcgccatgatgacatcacaaacagaggacagtaaacctgctaaatggctcaaaggt
 cacggtgggtgaaaagttggtggttagtggttctgttaaacgcacgtagagcaattcaaaga
 gcacaatctatttttaggtggacgttataaatcgtcttatattaccgaacaaatgatgaga
 gttgccaacacgtgagcttaacttccaatcagatgcggttaataactgggacatcaacgca
 caaaaagggaacgccttctaaaggatgtttccaaatgattgaaccatcttttagagcatat
 50 gctaaacacaggacacggaacatcttaaatccaaactgacgaagctatatctgctatgcgt
 tacattgtaggttaagtgggttccattatggggagttggagaagtgcatttaaacgtgct
 ggagattatgcttatgctacaggcgggttattaacactgctggattatataatttgga
 gaagatggataccctgagatagtaatccctacagatccaagcagacaatcagatgcgatg
 aaattgttacatcttgcgtcaagtaaaaattagtggaataaacagaaataaacgacctaac
 55 caattacgtacacctaatgttactagtaatacagttgataatgcagaattactactacaa
 atgatagaaaaatcaacagaaacaaataaacgtgttaattggaatagcacgaagtaataaa
 actattgaaaaacaaccgaaaggttttccagaacgcgatgtaagttaggcacaaggttca
 aggttaagactcgctgcttatagccaggaggtttataa

Sequence 2906

MKDMGVQRSISEIKRSFKGLNADLKLSNNNFKYSEKSLNSYKLRTRELSQAVKESKANVA
 ALKAKYQEAARASGVNSKKAQRLRQEYSRQADNLNLYQNELDQTRGKYREMIASVSSVG
 RLGOAFSEIGPKIRSIGDSMKSVGRNMSLHVTAPIAAGFGAAVKKSIDFDDTMRKVKATS
 5 GATGDEFNQLRTKALQMGRDTKFTASESAEAMNYMALAGWDTKDMLKGVGGVMDLAAASG
 EDLASVSDIVTDNLTAFGMKAKDSTHFADVLAQTSSKANTDVRGLGDAFKYAAPVAGALG
 YTVEDTSIAIGLSNAGIKGEKAGTALRTMFTNLSKPTKAMKDEMDKLGISITDSNGEML
 PMRDVLDQLRGKMGGLSKDQAAAAASTIFGKEAMSGALAVINASDEDYKKLTKSIDGSKG
 ASKRMAKEMEGGIGGAMRKMSAIESLAISLGDALAPMLYKAAKWITSLANKFSNLPTGV
 10 QKTIADVGLLAAAIGPLLMVFGVMASITIGTAITVLGSLMTSMRTLSFSLKTSAAATGIWN
 GVTATARGIANGRYAVAALTSQTIQAMKTKIAAAATTAWTTVTKGATLATKGLGLAIR
 FMTGPVGIVITAIYGLLVAGLIHLWKTNSFRNAVIGIWNISIKNAIAIFGFIKPYIINIW
 NAIKNSTIAIWNNAIKRSVAVIWNNAIKFAVQHPHQALKNVLSALWNGMKNAAIKIWTALKN
 GVIAIIKAYVAQVKFNINLIKRIVVTFNAIKSFSIKVWTALKNGVLGIVRALRKGVLSV
 15 FNALKKGIIVIFNAVKNAAVKIWTAIKKSUVNKAALWGVKNTWNALKKGTIGIFKAVG
 SFMSSKWNISIKKGTVNKAALWGVKGAWSLKKGTHNTMNAVGGFMSKKWNGIKSTTVS
 IVNGMKSVMGMTMNMRDGIKTVTGKIGNLFGGMVKGKGLNGLIKGVNWVADKLGMDK
 IPKIKLSTGTQSTHTQSYITNGKINKGTMATVGDGRGNGPGGFRNEMIRYPNGRMALTP
 NKDTTTTFLPKGSSVYSGAQTHALNSNGYDTKKKKLPKFSKGTKKKDGILDVSSGVKND
 20 VNKVKDIGGKARDIGGTTFDKAKDIGTKALDKAKDVSSSTVIKIGDVFVDYVGHPMKLVNK
 VFEKVGFNLDPMKNAPLPFDLMTAMIKKLKNGIKDFFNEGLDSAGGGDGSSFTKFPITTG
 YYPNGGAPGYSFNGGAHFGIDYAPYGTINATNDGNVKAHNLGGGLVARLLTGQFTLF
 FMHLSKILKQKIKAGEPMAKTGNSGQWTTGPHVHFQVERGRHDDITNRGTVNPAKWLG
 HGGGKVGSGSVNARRAIQRAQSILGGRYKSSYTEQMMRVAKRESNFQSDAVNNWDINA
 25 QKGTPSKGMFQMIEPSFRAYAKPGHGNILNPTDEAISAMRYIVGKWVPIMGSWRSFAFKRA
 GDYAYATGGVINTAGLYNLAEDGYPEIIVPTDPSRQSDAMKLLHLAASKISGNNRNKRPN
 QLRTPNVTSNTVDNAELLLQMIENQQKQINVLMELIARSNKTIEKQPKGFSERDVSQAQGS
 RLRLAAYSQGGGL*

30 Sequence 2907

Contig_0497_pos_4846_3491

is similar to (with p-value 9.0e-50)

>gp:gp|Y10528|PACIOAB_1 P.aeruginosa cioA and cioB genes. N
 ID: g2208963.

35 atggattcagtagaataagctcgattgttgacgggtatgacacttgacagtgcataatcata
 tttgcaacaattgggtgttggtatgccacttatgtttgcaatagcagagtttattggcata
 aagaaaaatgatgcgaattatatcacattggctaaaagatggctgaaaggctacacgata
 accgttgctgttgagtcgttacagggtactattatgtgtttacaactttcacttgctctgg
 ccaacatttatgaaaatggcggtcatgttatcgcatcaccactattcatggaaactttt
 40 ccattcttttttgaaagcaatctttttaagattttattatacacttgggaaagatttaaa
 aaataaatggacacattttttcatatctatacccggtattataggaggttcattctcagca
 ttcttttatatacatcagtcatttcatttatgaatacccccagctggttttgaaattaaaaat
 ggtcgtatggttaaatgttcagccattagaagcaatgtttaattcatcgtttatggttcgt
 gctttacatgtatgttgcaactgcaggtatgacgatggcggtttatattagcagccatcgca
 45 gcgttttaattattacgtcataatcatacagaagatagaacataaccatacaaaaagctctt
 aatttaagcatgattgttggtatcatcaataccggtcttttcgatgattgcaggagattta
 tctgctaaaatttttacacaaaagttcaaccagataagcttgacgcatatgaatggcattat
 gatacgcaatctcatgcgaatttagttcttttcggtgtattaaatgaaaaaacacatgaa
 gtttcaggagcactagagattcccggtacttagtttttttagcagataaatagctttaat
 50 acaaaaagttaaagggttttaaatgaatttctctaaaaatgaattacctcctatgatagtgc
 tactttttcgatttgatggtatcaatgggaatattctgttttatcatttcaggactatat
 atgttatttttaattgttaaaaagtacgtaaaatatgtaactagtaatatgatgtttatac
 gctattctgttaactggctcctgcatcaatgttagcgatagaatttggttggttcttaact
 gaaatgggtcgtcaaccgtggttatatacagagatatatgcgcgtatcagaagctgctacc
 55 caggcggggggaatcacattagtaacgacgctattcggtttactatatcttttattacta
 gttacttcggcttatgtactaattagaatgtttaaaaatcaacctgcttataaagatgta
 gaaaaagtgattaaagagagaggtgaaacaaaatga

Sequence 2908

MDSVEIARLLTGMTLAVHIIIFATIGVGMPLMFAIAEFIGIKKNDANYITLAKRWSKGYTI
 TVAVGVVVTGTIIIGLQLSLVWPTFMKMGGHVIALPLFMETFAFFFEAIFLSIYLYTWERFK
 NKWTHFFISIPVIIIGGSFAFFITSVNSFMNTPAGEFIKNGRMVNVQPLEAMFNSSFMVR
 ALHVVATAGMTMAFILAAIAAFKLLRHNHTEDRTYHTKALNLSMIVGFINTVLSMIAGDL
 5 SAKFLHKVQFDKLAAYEWHYDTQSHANLVLFGLVLEKTHEVSGALEIPGLLSFLADNSFN
 TKVKGLNEFPKNELPPMIVHYFFDLMVSMGIFCFIISGLYMLFLIVKKLRKYVTSNMMLY
 AILLTGPSMLAIEFGWFLTEMGRQPWIIRGYMRVSEATQAGGITLVTTFLGLLYLLLL
 VTSAYVLIRMFKNQPAYKDVEKVIKERGETK*

10 Sequence 2909

Contig_0499_pos_7860_8597

is similar to (with p-value 4.0e-71)

>sp:sp|Q46807|ARCL_ECOLI_CARBAMATE_KINASE-LIKE_PROTEIN_1. >
 gp:gp|U28375|ECU28375_24 Escherichia coli K-12 genome; appro
 15 ximately 64 to 65 minutes. NID: g887800. >gp:gp|AE000370|AE0
 00370_9 Escherichia coli K-12 MG1655 section 260 of 400 of t
 he complete genome. NID: g2367170.

gtgaatcatttgaatttatttttgggaaaaaagcacattcttaaagatattacattctca
 ttacctataaatgggtgaaatcattgggtattgtagacctaattgggtccggtaaatcatct
 20 tttctcaaaagcattttatcggaagaatttaaagcgaccggtaaacatacattatatgatcac
 cctatacacacaaacttagatatattacatatatccacaaaaagctcatattgattta
 gattttcctataaaaagtgatcaagtgatactatcaggatggtatgaagacattggatgg
 tttaaaaaagcaagtgtagttgaaaaaacgaaacttaaccaattacttaattgatttagaa
 ctgcacatatatcgccatcgacaaattgcagaattaagtgggtggacaattgcaacgtgtt
 25 cttgttgcaagagcactttatgagtaaatagtgtatatttattggttagacgaacctttgtt
 ggcatacgacattttatagcgaacaactttattatgaagaaaattaaacatttaagacatatg
 ggtaaatgtgattttaattgttcatcatgacttatcaaaagcagatcaatattttgaccgc
 attttattattaaatcaatcggttgcagtttttagggccaacaaaagaagcattatcatct
 gaacgattaaatgcaactttcattaattacaaagatgattcgcttttaacactatcctca
 30 caagggaagtacgaattag

Sequence 2910

VNHLNLFGLGKKHILKDITFSLPINGEIIIGIVGPNAGAKSSLLKAFIGEFKATGKHTLYDH
 PIHTQLRYITYIPQKAHIDLDFPIKVDQVILSGCYEDIGWFKKASVVEKTKLNQLNDLE
 35 LDHIRHRQIAELSGGQLQRVLVARALMSNSDIYCLDEPFVIGIDIYSEQLIMKKIKHLRHM
 GKLILIVHHDLSKADQYFDRILLNQLSLQFLGPTKEALSSERLNATFINYKDDSLTLSS
 QGSTN*

Sequence 2911

40 Contig_0499_pos_8598_9434

>gp:gp|X99127|SEABCTS_1 S.epidermidis gene encoding ABC tra
 nsport system. NID: g1617427.

atgtagatttcattaaccatttgcttagttatcaatttttaaatcggtgcattaatcaca
 tctatttttagttgggattgtatgtggaacgatgggttagcattattggtttacgtggtctt
 45 tctttaatgggtgatgccatgagtcagtcgtgtttaccaggtgttgctttatctttctta
 ttttaattccaatggtttatcggggcactttgtaacgggaatgcttgcaagtttggttatt
 ggtttttactttcaaacagtaaaaacaaaccagatgctgcaataggaataagtttcact
 gcattcctagcatctggcgtcataattatttagtttaataatagtagacaacagatttatat
 cacattttattttggcaattttattagcaattacacatcaatcattttggacaacaattgtc
 50 attactgtactgggtattttacttattattatcttttatagaccttaattgatttcaaca
 tttgatgcaacttttagtcgtatgagcgggctgaacacaacattaattcactactttgtc
 atgttattactcgcaacttgtaactgttgcgagcatacaaacagttggaattatccttgta
 gtggtttactaatcactccagcttctacagcttttttaatacagtaaaacaacttttcgcc
 atgatggtaattgcaagcataatcagcgtgataagttcgattatcggtctatatttttagt
 55 tatatatataatattccaagtggagcaactattgtaattctgtacctttatgatttatatt
 gtaacgctatcaattactagaattaaaaataaacaacaaaaggagcgctttaacgtga

Sequence 2912

MLDFINHLLSYQFLNRLITSILVGIVCGTMGSIIVLRGLSLMGDAMSHAVLPGVALSFL

FNIPMFIGALVTGMLASLFIGFITSNSKTKPOAAIGISFTAFLASGVIIISLINSTTDLY
HILFGNLLAITHQSFWTTIVITVLVILLIIIFRPLMISTFDATFSRMSGLNTTLIHVYFV
MLLLALVTVASIQTVGIILVVALLITPASTAFLISKQLYAMMVIASIIISVISSIIGLYFS
YIYNIPSGATIVICTFMIYIVTSLITRIKNKQKRSALT*

5

Sequence 2913

Contig_0499_pos_0_864

>gp:gp|X99127|SEABCTS_2 S.epidermidis gene encoding ABC tra
nsport system. NID: gl617427.

10 gtgagtgaatggctaaaattgtagtagcttttaggtgaaacgcttttaggaaaatcacca
caagaacaacttgaattagtaaaaaatacagctaaatccctagtaggattaattactaaa
ggcacgaaattgtgattagtcacggtaatggaccacaagtaggaagtattaaccttggt
ctgaattatgcagctgaacacgcatcaaggtcctgcttttccatttgctgaatgtggcgct
atgagtcgaagcctacatcggtcatcaacttcaagaaagtttacaataaacttcattca
15 atgggcatagataagcaagttgtcacactagttacccaagtagaagttgatgaaggcgat
ccagcttttaatagtcgaagtaaacccatcggtctgttctacactaaagaagaagcaaat
cgtattcaacaggaaaaaggttatcaattttagaagatgctggctcgaggttacccgtcgc
gttgtagcatcaccacaaccaatatctattatcgaactggaaagtattaaaactctagta
gaaaatgacacactcgtcatcgctgcaggtggaggtggatataccagtcattcgcgaacag
20 catgtagcttttaaggtatagatgcccgtcatcgataaaagacaaaacaagtgcattatta
gggtgcatattcactgtgatcaactcattattttaacagcgattgattatgtttatc
aactatcatactgaccaacaacaagcacttaaaacaacaatatagatacgcttaaaaca
tatattgaagaagaacaatttgccaaaggcagcatgctacctaataatcgaatctgccatc
tcctttattgaaaataatcctaac

25

Sequence 2914

VSEMAKIVVALGGNALGKSPQEQLELVKNTAKSLVGLITKGHEIVISHGNGPQVGSINLG
LNYAAEHQDQGPAPFPFAECGAMSQAYIGYQLQESLQNELHSMGIDKQVVTLVTVQEVDEGD
PAFNPSKPIGLFYTKEEANRIQQEKGYQFVEDAGRGYRRVVPSPQPIISIIIELESIKTLV
30 ENDTLVIAAGGGGIPVIREQHDSFKGIDAVIDKDKTSALLGADIHCDQLIILTAIDYVYI
NYHTDQQQALKTTNIDTLKTYIEEEQFAKGSMLPKIESAISFIENPN

Sequence 2915

Contig_0500_pos_3850_3113

35 is similar to (with p-value 2.0e-97)

>sp:sp|P36839|ARGD_BACSU ACETYLORNITHINE AMINOTRANSFERASE (EC 2.6.1.11) (ACOAT). >pir:pir|S38431|S38431 N-acetylornithine aminotransferase - Bacillus subtilis >gp:gp|Z26919|BSCITB
O_4 B.subtilis (168) DNA for argC-F citrulline biosynthetic
40 operon. NID: g408113. >gp:gp|Z99109|BSUB0006_199 Bacillus subtilis complete genome (section 6 of 21): from 999501 to 1209940. NID: g2633260. >gp:gp|Z99110|BSUB0007_4 Bacillus subtilis complete genome (section 7 of 21): from 1194391 to 1411140. NID: g2633472.

45 atgaaaaacatcatcgtaattaagctcggtggtatagctatagaaaaatttaaacgacgca
tttatacaacaaattaatgcttggcaccttgaaaacaaaaataattattgttcacggt
ggcgccaagtcacagtaatttatttaactaaaaacaatcattcaactattaaaattgat
ggcatgagagtaactgctaaaaatgatttacctatcatatatgatgctttaattaacata
gttgccatcaacttttagaaagacttaagaatctaatttagaatttttcaatttaa
50 gaaaagataaaagaacttgtaagcgccgaatttttaataaaaaatatctatggttacggt
ggtaaagttaaagaaatcaacacgatgctattagaaaaatgttatcacgcgacataata
ccaattatcactagtttgggtgtaaagtagcaaggggagtagtcttaattgttaattgctgat
catctcgccaccgcaattgctaagaaactaaaagtcgaaaaattagtatatatgactgat
gtccttggtgtaattgaaaaagataaaaacgcttgctactctacaattaatgaagcaaaa
55 acaaaaattgaaaataaaaataattacaggcggaatgatacctaataattgagagtgaatc
caaacattagaatctggtggtgaatcgattttaattgcaataatttacaaaaaggaaca
atcataagggtgattaa

Sequence 2916

MKNIIIVIKLGGIAIENLNDAFIQQINAWHLENKKIIIVHGGGQVISNLLTKNNHSTIKID
 GMRVTAKNLDPPIYDALINIVGHQLLERLKESNLEFFQFKEKIKELVSAEFLNKNIIYGYV
 GKVKEINTMLEKMLSRDIIPIITSIGVNEQGEYLVNADHLATAIAKKLKVEKLVYMTD
 VPGVIEKDKTLATLTINEAKTKIENKIITGGMIPKIESAIQTLESGVESILIANNLQKGT

5 IIRGD*

Sequence 2917

Contig_0500_pos_3110_1983

is similar to (with p-value 1.0e-34)

10 >sp:sp|P36840|ARGB_BACSU ACETYLG LUTAMATE KINASE (EC 2.7.2.8
) (NAG KINASE) (AGK) (N-ACETYLG LUTAMATE 5-PHOSPHOTRANSFERASE
). >pir:pir|S38430|S38430 N-acetylglutamate 5-phosphotransfe
 15 rase - Bacillus subtilis >gp:gp|Z26919|BSCITBO_3 B.subtilis
 (168) DNA for argC-F citrulline biosynthetic operon. NID: g4
 08113. >gp:gp|Z99109|BSUB0006_198 Bacillus subtilis complete
 genome (section 6 of 21): from 999501 to 1209940. NID: g263
 3260. >gp:gp|Z99110|BSUB0007_3 Bacillus subtilis complete ge
 20 nome (section 7 of 21): from 1194391 to 1411140. NID: g26334
 72.
 atgagttatctttttaataattacaagcgtgacaatatagagtttggtgatgctaataca
 aatgaattaattgataaagataataatgtctacctagatttttcgtcaggtataggtgtg
 acaaatctgggttttaatatggaattttaccaagcagtttataatcaactgaatttaata
 tggcattccaccaattttatacctaagtagtatccaagaggaagtggctcaaaaattaatt
 ggtcaacgagattatcttagctttcttttgtaatagcggaaacagaagcgaatgaggcagct
 25 atcaactcgcacgtaaagctactggttaagtcggaaattattgcttttaaaaagctcttt
 ccaggcagaacgtacggcgcaatgtctgcaacaggacagaagaaaattacagatcaattt
 ggtccggttggtccctggattcaatttgctatttttaattgatttttaattcatttaaatca
 ttaacttcaataataactgctgctgtaattatagaaataattcaaggtgaatcaggagta
 ctacctgctgattctttatttatgaagcaattaaatgagtattgtaaacaaaaagatatc
 30 ctataattgtagacgaggttcaaacgggcataggtagaacccggttaagttatatgctcat
 gaacattatcaattgtctccagatatcatcacattagctaaaggattaggtaatggcctt
 cctattggagcaattgttaggcaaaaaaatttaggtcatgcatgttggtacggttctcat
 ggtacaacattcgggtggaaatagattatcattggctgctgcaaaccaaacgctttctatc
 attaatgatgctgatttgctgaatgatgttcaatcaaaaggggcaatttcttattgaaaac
 35 ttaagaaaaagtttagtaataaaaagaaatgtaattgaagtacgtggtgtaggtttaatg
 gtaggaatagaggtcactaatgatcctagtcagtagtgcgagaagctaaacgtatgggg
 ttaatcatttttaacagctggtaaaaatgtgattaggttattaccgccattgaccatcact
 aaaaaacaattagaaaaaggtatagaaatattaactgaaatcatttga

40 Sequence 2918

MSYLFNNYKRDNIEFVDANQNELIDKDNVYLDFFSSGIGVTNLGFNMEIYQAVYNQLNLI
 WHSPNLYLSSIQEVAQKLIGQRDYLAFFCNSGTEANEAIAIKLARKATGKSEIIAFKKSF
 HGRTYGAMSATGQKKITDQFGPVVPGFKFAIFNDFNSFKSLTSNNTAAVIEIIQGESGV
 LPADSLFMKQLNEYCKQKDILIIIVDEVQTGIGRTGKLYAHEHYQLSPDIITLAKGLGNGL
 45 FIGAMLGKKNLGHAFGYGSHGTTFGGNRLSLAAANQTLISIINDADLLNDVQSKGQFLIEN
 LRKSLVNKRNVIEVRGVGLMVGIEVTNDPSQVVREAKRMGLIILTAGKNVIRLLPPLTIT
 KKQLEKGIIEILTEII*

Sequence 2919

50 Contig_0502_pos_13693_14094

is similar to (with p-value 7.0e-44)

>sp:sp|P22806|BIOF_BACSH 8-AMINO-7-OXONONANOATE SYNTHASE (E
 C 2.3.1.47) (7-KETO-8-AMINO- PELARGONIC ACID SYNTHETASE) (7-
 KAP SYNTHETASE) (L-ALANINE--PIMELYL COA LIGASE). >pir:pir|JQ
 55 0512|JQ0512 8-amino-7-oxononanoate synthase (EC 2.3.1.47) -
 Bacillus sphaericus >gp:gp|M29291|BACBIOXWF_3 B.sphaericus b
 ioXWF operon genes, complete cds. NID: g142592.
 gtgcacctgattgtgctggcgtacctgggttccactgggtttacctgggtctgctggcgtac
 ctggttccgctgggtttacctgggttccgctgggtttacctgggttccgctgggtttacctggtt

ccgctgggtttacctgggttctgctgggtttacctgggtccgctgggtttacctgggttctgctg
 gcgtacctgggtccgctgggtttacctgggtccgctgggtttacctgggttctgctgggtttac
 ctgggttctgctggcgtacctgggtccgctgggtttacctgggttctgctggcgtacctgggt
 ccgctgggtttacctgggtccgctgggtttacctgggtccgctgggtttacctgggttctgctg
 5 ttgttgaccatactcaacaatttcgtcaacagggtgttag

Sequence 2920

VH1:IVLAYLVPLVYLVLLAYLVPLVYLVPLVYLVPLVYLVLLVYLVPLVYLVLL
 AYLVPVYLVPLVYLVLLVYLVLLAYLVPLVYLVLLAYLVPLVYLVPLVYLVPLVYLVLL
 10 LLDHTQQFRQQVV*

Sequence 2921

Contig_0502_pos_15619_13529

is similar to (with p-value 3.0e-21)

15 >sp:sp|P22818|BIOD_BACSH DETHIOBIOTIN SYNTHETASE (EC 6.3.3.
 3) (DETHIOBIOTIN SYNTHASE) (DTB SYNTHETASE) (DTBS). >pir:pir
 |JQ0506|JQ0506 dethiobiotin synthase (EC 6.3.3.3) - Bacillus
 sphaericus >gp:gp|M29292|BACBIODAYB_1 Bacillus sphaericus I
 F03525 bioDAYB operon encoding dethiobiotin synthase (bioD),
 20 adenosylmethionine-8-amino-7-oxononanoate aminotransferase
 (bioA), biotin synthase (bioY) and bioB genes, complete cds.
 NID: gl42587.

gtggatgagatcggttcattatggtggcgaagaaatcaagccaggccataaggatgaattt
 gatccaaacgcaccgaaaggtagccaagaggacgttccaggtaaaccaggaggttaaaaat
 25 cctgatacaggcgaagtagtcacaccaccagtgatgatgtgacaaaatattggtccagtt
 gatggagattcgattacgtcaacggaagaaattccattcgacaagaaacgtgaattcaat
 cctgatttaaaaccagggtgaagagcgtgttaaacagaaagggtgaaccaggaacaaaaca
 attataaacaccaacaactaagaacccattaacaggggaaaaagttggcgaaggtg.acca
 acagaaaaaataacaaaacaaccagtagatgaaatcacagaatatggtggcgaagaaatc
 30 aagccaggccataaagatgaatttgatccaaatgcaccgaaaggttagccaagaggacgtt
 ccaggtaagccaggaggttaaaaaatcctgatacaggcgaagtagtcacaccaccagtggt
 gatgtgacaaaatatggtccagttgatggagatccgattacgtcaacggaagaaattccg
 ttgataaaaaacgcgaatttaattccaaacttagcgccaggtagacagagaaagtcgttcaa
 aaagggtgaaccaggaacaaaaacaattacaacaccaacaactaagaacccattaacaggg
 35 gaaaaagttggtgaaggtgaaccaacagaaaaaataacaaaacaaccagtgatgagatc
 gttcattatggtggcgaagaaatcaagccaggccataaggatgaatttgatccaaacgca
 ccgaaaggtagccaagaggacgttccaggtaaacaggaggttaaaaaatcctgatacaggc
 gaagtagtcacaccaccagtggtgatgtgacaaaatatggtccagttgatggagattcg
 attacgtcaacggaagaaattccggttgataaaaaacgcgaatttgatccaaacttagcg
 40 ccaggtagacagaaagtcgttcaaaaagggtgaaccaggaacaaaaacaattacaacgcca
 acaactaagaacccattaacaggagaaaaagttggcgaaggtgaaccaacagaaaaata
 acaaaaacaaccagtggtgagattgttcattatggtggtgaacaaataccacaaggtcat
 aaagatgaatttgatccaaatgcacctgtagatagtaaaactgaagttccaggtaaacca
 ggagttaaaaatcctgatacagggtgaagttgttaccaccaccagtggtgatgtgacaaaa
 45 tatggtccgaaaagttggttaattccaatcacatcaacggaagagattccatttgataagaaa
 cgtgtattttaattcctgattttaaaaccagggtgaagagcgcgttaaacaaaaaggtgaacca
 ggtaaaaaacaattacaacaccaatattagtttaattcattacaggagaaaaagttggc
 gaaggttaataacagaaaaagtcactaaacaacctgttgacgaaattgttgagtattggt
 ccaacaaaagcagaaccagggtaaaccagcggaaccagggtaaaccagcggaaccagggtaaa
 50 ccagcggaaccagggtacgccagcagaaccagggtaaaccagcggaaccagggtacgccagca
 gaaccagggtaaaccagcagaaccagggtaaaccagcggaaccagggtaaaccagcggaacca
 ggtacgccagcagaaccagggtaaaccagcggaaccagggtaaaccagcagaaccagggtaaa
 ccagcggaaccagggttaaccagcggaaccagggtaaaccagcggaaccagggtaaaccagcg
 gaaccagggtacgccagcagaaccagggttaaccaggtgaaccagggtacgccagcacatca
 55 ggtgcaccagaacaacaaatagatcaatgcattcaacagataataaaaaatcaattacct
 gatacagggtgaaaatcgtaagtaattgagggaaactttagtcggatctctattagcaatt
 gtcggatcattgttcatatttggtcgctcgtaaaaaaggtaatgaaaaataa

Sequence 2922

VDEIVHYGGEEIKPGHKDEFDPNAPKGSQEDVPGKPGVKNPDTGEVVTPPVDDVTYGPV
 DGDSITSTEEIPFDKKREFNPNLKPGEERVVKQKGEPTTKITTTPTTKNPLTGEKVGEGETE
 TEKITKQPVDEITEYGGEEIKPGHKDEFDPNAPKGSQEDVPGKPGVKNPDTGEVVTPPVDD
 5 DVTKYGPVDGDPITSTEEIPFDKKREFNPNLAPGTEKVVQKGEPTTKITTTPTTKNPLT
 EKVGEGETEKEITKQPVDEIVHYGGEEIKPGHKDEFDPNAPKGSQEDVPGKPGVKNPDTG
 EVVTPPVDDVTYGPVDGDSITSTEEIPFDKKREFNPNLAPGTEKVVQKGEPTTKITTTPT
 TTKNPLTGEKVGEGETEKEITKQPVDEIVHYGGEQIPQGHKDEFDPNAPVDSKTEVPGKP
 GVKNPDTGEVVTPPVDDVTYGPVGNPITSTEEIPFDKKRVFNPNLKPGEERVVKQKGEPT
 10 GTKITTTPIVLNPNITGEKVGEGETEKEVTKQPVDEIVEYGPTEKPAEPGKPAEPGKPAEPGK
 PAEPGTPAEPGKPAEPGTPAEPGKPAEPGKPAEPGKPAEPGTPAEPGKPAEPGKPAEPGK
 PAEPGKPAEPGKPAEPGKPAEPGTPAEPGKPAEPGTPAEPGKPAEPGKPAEPGKPAEPGK
 DTGENRQANEGTLVGSLLAIVGSLFIFGRRKKGNEK*

Sequence 2923

15 Contig_0502_pos_12781_12110
 is similar to (with p-value 2.0e-34)
 >gp:gp|U67763|PBU67763_1 Plasmodium berghei thrombospondin
 related adhesion protein (PbTRAP) gene, complete cds. NID: g
 1906578.
 20 atgaatatatttgaactggaacgaatactgacatcggtaaaacgtatgtcactaaatat
 ctctataaggcattaagaacgagggatattcggtatgtattttcaaaccctttcaaact
 gaagaaattggtggaggttagataccagatttagaaattataaaaacgaatgagattta
 gactatgacgttacgtctctttacacattcaaagatccagtttcaccacatttagcattc
 aaaattgaaaggcatcagcaattgaacatcaacaatgatagacaaactcgaatcacta
 25 gaagcacaattcgatatgattctcattgaaggtgcaggtggtattgcagtgcctatctat
 gaatacagtgaccatttttatatgacaacagatttaattaaagacacatcggtttcatt
 gtgagtgctctaccttcaaagttaggtgagattaatgatgccatcggttcaccagaaatat
 attgatcatcaagaacttccccgaatgtgttaatcatgaataactatactgatagtgt
 attgaacaggataatttacataccattgaaaaattaatacacaagctctgtttatacgttg
 30 ggtcatcaagcgactcaagaaagcttttcgaagcatttatacaacgaattataggagga
 tccaatggctaa

Sequence 2924

35 MNIFVTGTNTDIGKTYVTKYLYKALRTRGYRVCIFKPFQTEEIGGGRYPDLEIYKNECDL
 DYDVTSLYTFKDPVSPHLAFKIERHQQLNHQTMIDKLESLEAQFOMILIEGAGGIAVPYI
 EYSDHFYMTDLIKDTSDFIVSVLPSKLGAINDAIVHQKYIDHQELPPNVLIMNNYTDSA
 IEQDNLHTIEKLIHKSIVYTLGHQATQESFSEAFIQRIGGSNG*

Sequence 2925

40 Contig_0502_pos_10751_9618
 is similar to (with p-value 8.0e-36)
 >gp:gp|U67763|PBU67763_1 Plasmodium berghei thrombospondin
 related adhesion protein (PbTRAP) gene, complete cds. NID: g
 1906578.
 45 atggacattaaagcacagttaaaacagattcaagataaaggtttatatagagagcttcag
 ccgattcagtcgtagaaaaacaatatatttatatcaatgaccaatcttatattaatttt
 acttcgaacgattatctcggtataggacaagttgaatatcaacctcaaaatttcttagat
 tttataaagacatatagtagtccatctatcaagttctagatttagtgagtggaattcagtt
 50 gtttatcagcaattagaacaggaaattagcgagcattttaattttgaagacgccttaatt
 tttaatagtggttacgatgcgaatttgccggtatttaatatattttaaaaataataatatt
 gttatttttctgatcaacagaatcatgccagtataatagacggtattaaattaagtggg
 ttatcaaaagtgtatttatcaacatttaaaactatgatgacctggaagtcatttagcacgg
 cacaccaatccagatgttcaaaaagtaattgtctctgatagtgtgttttctactaatggc
 actaaagcagatattaataggctagtacatctcaagcaacgttacaatgcgattttaatt
 55 attgacgcattctcatagtttaggattaaatctcttgagtatcatgcagacattgacata
 gttacttcaagtttatctaaagcggtggggagcccatgggtggcgtcatattcagttcaaaa
 gatataaaagatttaattcattaataaagggtcgttcgcttatctactcgagtagttacct
 agctatcatattgtattttattcaagtgagcttacaacatgtgattgaagatacataaga
 cgagagagagttgaatgtacttagtgaatatatttaaccacattcatggaattatttccc

gatcaaccattatctaatcacctatcaaaaatatcgtttgatagtttggttcagca
 caagcacaatacgacatgttatttgaacatggtatatttgcagttatttaaggtatcca
 acagtgtcacagctaacattaagaatttcattatcctattttcatgacacagatgatatt
 gatcgacttttcaatgtaatgaaacaatacgatgaaggtgatagctatgtatag

5

Sequence 2926

MDIKAQLKQIQDKGLYRELQPIQSVEKQYIYINDQSYINFSTNDYLGIGQVEYQPQNFLD
 FIKTYSIHLSSRLVSGNSVYQQLQEISEHFNFDALIFNSGYDANLAVFNIFKNNNI
 VIFSDQONHASIIDGIKLSGLSKVIYQHLNYDDLESHLARHTNPDVQKVIVSDSVFSTNG
 TKADINRLVHLKQRYNAILIIDASHSLGLNLFYHADIDIVTSSLSKAWGAHGGVIFSSK
 DIKDLIINKGRSLIYSSSLPSYHLYFIQVSLQHVIEDTYRREKLNVLSEYFNHQFMELFP
 DQPLSNTPIKNIVCDSLASAQAQYDMLFEHGIFVSYLRYPTVSQLTLRISLSYFHDITDDI
 DRLFNVMKQYDEGDSYV*

15

Sequence 2927

Contig_0509_pos_6421_7818

>sp:sp|P13408|UHPT_ECOLI HEXOSE PHOSPHATE TRANSPORT PROTEIN
 . >pir:pir|A30395|MMECHP hexose phosphate transport protein
 uhpT - Escherichia coli >gp:gp|M17102|ECOUHP_5 E.coli uhp op
 eron encoding UhpA, UhpB, UhpC, and UhpT protein, (encoding
 20 hexose phosphate transport protein), complete cds, and an il
 vBN operon encoded protein, 3' end. NID: g148110. >gp:gp|M89
 479|ECOUHPABCT_4 Escherichia coli uhpABCT operon encoding he
 xosephosphate utilization protein (uhpA) gene, complete cds,
 25 and hexosephosphate transport protein (uhpB, uhpC, uhpT) ge
 nes, complete cds. NID: g148116. >gp:gp|AE000444|AE000444_5
 Escherichia coli K-12 MG1655 section 334 of 400 of the compl
 ete genome. NID: g2367258.

30

35

40

45

50

atgataggggtgttagggatgaactttttgatattcataaaatgccaaacaaagggata
 ccattagctgtacaacgcaaattatggctcagaaactttatgcaagcggtttttgtgta
 ttctttgtttacatggcgatgtatttaattcgaacaattttaaagcggcacaac:gtta
 ttaaaagaagaatcggattaacaacattagaactaggttatataggattagcgttttagt
 attacttacgggttaggaaaaacaatactcgggtatttcggtgatggcgtaatacgaaa
 cgtattatttcttcttatttaataattatctgcgattacagtacttattatgggatttgta
 ttaagttatttcggttctgtgatggggctattaattgtattgtggggcctaacggtata
 tttcaatctgtgggtgggcctgcaagttactcaacgatttcaaggtgggcgcctcgaca
 aagcgcggtcggtatttaggcttttggaatacatcacataacattgggtggtgctattgct
 ggtggtgctgcactttggggcgcaataacattttccacggtaattgtggttggaatggtt
 atttttccttccgctcatcgctttaatcattgggattgtgacattatttattggtaaagat
 gatccagaggaattagggtggaatcggtgccgaagaaatttgggaagagcctatcgacca
 gaaaacattgattctcaaggtatgactaaatgggatattcttataaaatatatccttggga
 aatcctgtgatttggattttgtgtatctctaattgtttttgtatatatcggtgcgtattggt
 attgataactgggcaccgctatcgtatcagagcatttacattttaataaaggtgatgcg
 gtgaatactatttttactttgaaatagggtgcattagtagctagtttattgtggggctat
 atctcagatttataaaaggtcgctgcgtgcatgttagcgattggatgtatgtttatgatc
 acctttgttgactcttttataccaatgcaacaagcgtgacaatgggtcaatatttctcta
 tttgcattaggcgctttaatcttcggtccacagttaactcattgggtgtatctctgactggc
 tttgttctctaaaaatgcaattagtgctgctaacgggtatgacaggttcatttgcatactta
 ttcggggattcaatggctaaagtggtctggctgcaatcgctgatccaacacgtaatggt
 ttaaatatttttgggtatacgttgagtggttgacagatgtctttattgtattctatgta
 gctttattcttaggaatgatattatagccattgttgcttattacgaagaaaagasaatt
 agaaaattaaaaatttaa

Sequence 2928

MIGVLGMNFFDIHKMPNKGIPLAVQRKLWLRNFMQAFFVVFVYMAMYLRNNFKAAQPL
 LKEEIGLTTELGYIGLAFSITYGLGKTIILGYFVDGRNTRKRIISFLLILSAITVLIMGFV
 LSYFGSVMGLLIVLWGLNGIFQSVGGPASYSTISRWAPRTKRGYLGFWNTSHNIGGAIA
 GGVALWGANTFFHGNVGMFIFPSVIALIIGIVTLFIGKDDPEELGWNRAEEIWEEPIDQ
 ENIDSQGMTKWDIFKKYILGNPVIWILCISNVFVYIVRIGIDNWPPLYVSEHLHFNKGDA

VNTIFYFEIGALVASLLWGYISDLLKGRRAIVAIGCMFMITFVVLFTNATSVTMVNISL
FALGALIFGPQLLIGVSLTGFPKNAISVANGMTGSFAYLFGDSMAKVLAAIADFTRNG
LNIFGYT'LSGWTDVFFIVFYVALFLGMILLAIVAYYEEKKIRKLKI*

5 Sequence 2929

Contig_0517_pos_10178_9693

is similar to (with p-value 2.0e-26)

>sp:sp|P55978|GREA_HELPY TRANSCRIPTION ELONGATION FACTOR GR
EA (TRANSCRIPT CLEAVAGE FACTOR GREA). >gp:gp|AE000596|HPAE00
10 0596_20 Helicobacter pylori section 74 of 134 of the complet
e genome. NID: g2313982.

atgagatttatggaaaacaaaaacaatatcctatgactcaagaagggttatgagaaactt
gaacaagaattagaagaattaaaaacgggttaaaagacctgaggtagttgaaaaataaaa
gtagctcgttcatttggagacctatctgagaactctgaatatgatgctgctaaagatgaa
15 caaggctttattgaacaagatatacaacgtattgaacatatgattagaatgcgttaac
attgaagataacgggtgataacaatgtagttcaaatggtaaaacagttactttttattgaa
ttacctggagatgaagaagaaagttatcaaactcgttggttctgctgaagctgacgcattt
aaaggaaaaatttctaacgaatctccaatggcaaaagcactaatcggttaaaggattaaat
gatcaagtacgtgttccacttctaacgggtggcgaaatgaatgttaaaatcgttgaaatt
20 aaataa

Sequence 2930

MRFMENQKQYPMTQEGYEKLEQELEELKTVKRPEVVEKIKVARSFGLSENSEYDAAKDE
QGFIEQDIQRIEHMIRNALIIEDNGDNNVVQIGKTVTFIELPGDEEESYQIVGSAEADAF
25 KGKISNESPMALIGKGLNDQVRVPLPNGGEMNVKIVEIK*

Sequence 2931

Contig_0517_pos_9512_8913

is similar to (with p-value 1.0e-57)

>sp:sp|P24247|PFS_ECOLI PFS PROTEIN (P46). >pir:pir|S45227/
S45227 purine nucleoside phosphorylase homolog - Escherichia
30 coli >gp:gp|D26562|ECO82K_47 Escherichia coli genome, 2.4-4
.1 min region (110,917-193,643 bp from 0 min). NID: g473770.
>gp:gp|U70214|ECU70214_10 Escherichia coli chromosome minut
35 es 4-6. NID: g1552727. >gp:gp|AE000125|AE000125_6 Escherichi
a coli K-12 MG1655 section 15 of 400 of the complete genome.
NID: g1786348. >gp:gp|U24438|ECU24438_1 Escherichia coli MT
A/SAH nucleosidase gene, complete cds. NID: g2981266.

atgcgctatgcgtatgggtcagaaaagggtttaccggtatctaaaaaacatctcatcat
40 gctagaatatatattgggtcagccaacaaaatattcaacacgaggagatgcttttttatg
tcattctgacacaaacagtttagcacatacaaaatggaattgtaagtatcacatagt::ttt
gtaccacaaatatagaagacaagtgatatacggaaaaatcaaaagagatattggagtatt
ttacgtcaactatgtgaaaagaaaaggcgtagaataatagaagcagaagcatctaaagat
catattcatatgttagtttagtattccacctaataataggagtagtcctcatttgggtgat
45 ttaaaaggtaaaagcagtttaattgatttttgatagacatgctaatttaaaatatagatat
ggaaatagaaagtttgggtgtaaaggattttatgtagatacagtaggtagaaataaaaag
gtaattgaaaattacattcgtaataattacaagaagatatcggtgcagatcaaatttca
atggaagaatacctagatccctttacaggagaagaaattaaaaaaagacgaaaaaaatag

50 Sequence 2932

MRYAYGSEKVLPLSKKTSHHARIYIGOPTKIFNTRRCFFMSSDTNSLAHTKWNCXYHIVF
VPKYRRQVIYGKIKRDIGVILRQLCERKGVETIEAEASKDHIHMLVSIIPKLGVSFVGY
LKGKSSLMIFDRHANLKYRYGNRKFWCKGFYVDTVGRNKKVIENYIRNQLQEDIVADQIS
MEEYLDPTGEEIKRRKK*

55

Sequence 2933

Contig_0517_pos_4627_4100

is similar to (with p-value 1.0e-41)

>gp:gp|AF006000|AF006000_5 Bordetella pertussis D-3-phospho

glycerate dehydrogenase homolog (serA) and Brg1 (brg1) genes
, complete cds. NID: g2290988.

atggatcttactccaaatgaaatattataacttagttatatatcaattaggtgcgttaagt
ggcttttgtaaaatcaatcatgtaaaaatgatgcatgttaaacctcatggtgccctttat
5 caaatgggggctagaaataaaagaaattgcacatgcaattgctcaagcagtttttgatttc
gactcaaatctaatttttcgtcggcttagcgaatacattacttatttcggaagctgaatta
gtggggcttaaggtagcttcggaagtattttgctgaccgtcgttatgaagatgacggacaa
ttggtaaagtagaaaaaaaaccgatgccactatcactaatactgacgaagcaatccaacaa
gcattaaaaatggtttttgaaaaataaagttgtaagtaaaaatggaaaaatcatcgatttg
10 aaagctgatacaattttgtgttcacggagatggaaaacacgcattagaattttgttacgcaa
attagaaatgaattaatgaaagaaggcattgatattcaatccttatag

Sequence 2934

MDLTPNEIYNLVIYQLGALSGFCKINHVKMMHVKPHGALYQMGARNKEIAHAIAQAVFDF
15 DSNLIFVGLANTLLISEAELVGLKVASEVFADRRYEDDQLVSRKKTDATITNTDEAIQQ
ALKMVLNKKVSKNGKIIDLKADTICVHGDGKHALEFVTQIRNELMKEGIDIQSL*

Sequence 2935

Contig_0517_pos_2645_1959

20 is similar to (with p-value 2.0e-52)

>gp:gp|AF025380|AF025380_1 Salmonella typhimurium IS200 tra
nsposase (tnpA) gene, complete cds. NID: g2555163. >gp:gp|Y0
9990|STFLGLIS2_2 S.typhi flgL gene, gene encoding putative I
S200 transposase and gene encoding putative RNaseE-like prot
25 ein. NID: g2765044. >gp:gp|Y09991|STIS2T157_1 S.typhi encodi
ng putative IS200 transposase, 1575bp. NID: g2765048. >gp:gp
|Z54217|STISFLIBC_3 S.typhimurium fli[B,C] genes and inserti
on sequence IS200. NID: g1150641. >gp:gp|U44749|STU44749_1 S
almonella typhimurium putative IS200 transposase gene, compl
30 ete cds. NID: g1177216. >gp:gp|AF093749|AF093749_2 Salmonell
a typhimurium strain LT2 NADP+-linked malic enzyme (maeB), p
artial cds; insertion element IS200 transposase, complete cd
s; put operon, complete sequence; and unknown genes. NID: g3
885908. >gp:gp|L25848|STYIS200A_1 Salmonella typhimurium IS2
35 00 insertion sequence from SAR17, partial. NID: g439618.

atgataggaattattggagcaatggaagaagaagtgacgattttaagcgtaaattgaat
gatatgaatgaaataaattatgacgatgttaaattttatgttggcaagctaaaccacaaa
gaggtggttttaacacaaaagtggataggttaaagttaatgcttctatctcaacgactttg
40 ttaatagaaaaatttaattccagaagctgcattatactgacgaggtgcactagat
caaacacatatctattggagatatattagtgagtaatacatgtattatcatgatgcta
gctacacgcgtttggttatgaatatggacaaatacctcaaatgcctaaaacttatactact
gatcctactttgttgaaaaaacaatgcatgtattagaacaacaacaactgaatggtaaa
gtaggtatgattgttagtggtgatagttttataggtagctcagaacagcgacaaaaaatt
aagcaacaatttccagaagctatggctgtcgaaatggaggcaactgcaattgacgaaaca
45 tgttatcaatttaaaagtaccattttatcgtaactagagctgtttctgatttagcaaacggt
aaagccgatatttcttttgagaatttttagataaaagcagctttatcatctagttagaca
gtttcattattagtagaatcattataa

Sequence 2936

50 MIGIIGAMEEEVTILKRKLNDMNEINIAHVKEFYVGKLNHKEVVLTQSGIGKVNASISTTL
LIEKFNPVVINTGSAGALDQTLSDILVSNHVLYHDANATAFGYEYEQIPQMPKTYTT
DPTLLKKTMHVLEQQQLNGKVGMI VSGDSFIGSSEQRQKIKQQFPEAMAVEMEATAAQT
CYQFVVPFIVTRA VSDLANGKADISFEEFLDKAALSSSETVSLIVESL*

55 Sequence 2937

Contig_0518_pos_1926_964

is similar to (with p-value 1.0e-49)

>sp:sp|P45510|DHAK_CITFR DIHYDROXYACETONE KINASE (EC 2.7.1.
29) (GLYCERONE KINASE). >gp:gp|U09771|CFU09771_2 Citrobacter

freundii DSM 30040 cyclopropane fatty acid synthase (cfa) gene, partial cds, dihydroxyacetone kinase (dhaK), glycerol dehydrogenase (dhaD), transcriptional activator (dhaR), 1,3-propanediol dehydrogenase (dhaT), glycerol dehydratase (dhaB), glycerol dehydratase (dhaC) and glycerol dehydratase (dhaE) genes, complete cds. NID: g1229153.

atgaaaaagttaattcaagataaaaaacacaattttaaaagatatgcttgatggaattacagtttcaaacaacgatgttgaaagttgtatctgacactattgttgtagaaagcataaaaaa
 5 caatcaggtgttgactcgtttctggggcggcagtggaacatgaacctgcacacgcagga
 10 tttgtagcagaaggcatgctcgatgcagctgtatgtggagaaatcttcacttcacctaca
 cctgataaaatattagatgccattaaagctgtggacaatggtagcggcgttctacttgtt
 attaaaaactatgcaggagcgttatgaactttgaaatggctcaagaaatggctcaaatg
 gaagatatataaagttgaaagtgttattgtcagagatgatattgctatttctgacccggaa
 aaacgccgtggtgtagcgggtacagtattttgtgcataaatatgctggatacctagctgaa
 15 aaaggtgttgacttgatgaaatcaaatctaaagttgaggcacttttaccagatattaaa
 agtattggtatggcattaacgcctccaatggtgcccacaactggtaaaaacggtttcgat
 attgaagacaatcaaatgaaattggtatcggtattcacggtgaaaaaggtttacatcgt
 gaagacgtacaacctattaacgtgattgttgaaactttactcgatcaattatacaaaagaa
 attgagaaaaaaacccctttaatcgtaattggttaattggtatgggtggtacgccactatcagaa
 20 ttaaatatagttactaaatatctagatgaacaattcaatcagaatgatattggtgttaaa
 caatggttcgtagggtgactatatgacagcgttagacatgcaaggcttctctataactgta
 ctcccttcagtgaagaattgagtgaagctttagctgcacctacagcaagtaaatatttc
 taa

25 Sequence 2938
 MKKLIQDKNTILKMDLGITVSNNDVEVVSdTIVVRKHKKQSGVALVSGGGSGHEPAHAG
 FVAEGMLDAAVCGEIFTSPTPKILDIAIKAVDNGDGVLLVIKKNYAGDVMNFEMAQEMAQM
 EDIKVESVIVRDDIAISDPEKRRGVAGTVFVHKYAGYLAEGVALDEIKSKVEALLPDIK
 SIGMALTPPMVPTTGKNGFDIEDNQMEIGIGIHGEKGLHREDVQPINVIVERLLDQLYKE
 30 IEKKPLIVMVGMMGGTPLSELNIVTKYLDEQFNQNDIGVKQWFGDYMTALDMQGFSTIV
 LPFSEELSEALAAPTASKYF*

Sequence 2939
 Contig_0519_pos_4309_3827
 35 is similar to (with p-value 2.0e-25)
 >gp:gp|AF008183|AF008183_1 Populus balsamifera subsp. trichocarpa X Populus deltoides 4-coumarate:CoA ligase 2 (4CL2) mRNA, complete cds. NID: g2911796.

atgaataaaagtctattttgagcaatctcctcatTTTTGTAAAAATCACATTCATATTCCT
 40 gatggcttatcagaaaaatctagaagcagaagcggaaacgatataacaatttattagatgaa
 agagggccaatcgatattcaaatTTTAGGAATTGGAGAAAATGGTCACATTGGTTTAAAT
 gaaccagggaactgacttcaatagtgaacacacatgtggtgaacttaacagaaagcaccata
 aaagcaaatagtcgattTTTTGACAATGAAAAGGATGTTCTAGACAAGCAGTTTCAATG
 45 ggggtaaaaagtattttaaaagcaaaaaggattatcctactcgcatTTTGGTCCAAAGAAA
 aaagaggctataagtaaaactgttaaatgaacagggttaccgaagatgtacctgcgaccatt
 ttacacacacaccctaattgttgaagtttatgtagacgatgaagcagcgccagattgttta
 taa

Sequence 2940
 50 MNKVLFEQYPHFVKNHIHIPDGLSENLEAEAEERYNNLLDERGPIDIQILGIGENGHIGFN
 EPGTDFNSETHVVNLTSTIKANSRFFDNEKDVPRQAVSMGVKSILKAKRIILLAFGPKK
 KEAISKLLNEQVTEVPATILHTHPNVEVYVDDEAAPDCL*

Sequence 2941
 55 Contig_0519_pos_1586_288
 is similar to (with p-value 2.0e-22)
 >sp:sp|Q04802|NAG1 CANAL GLUCOSAMINE-6-PHOSPHATE ISOMERASE (EC 5.3.1.10) (GLUCOSAMINE-6-PHOSPHATE DEAMINASE). >pir:pir|A46652|A46652 glucosamine-6-phosphate isomerase (EC 5.3.1.1

0) - yeast (*Candida albicans*) >gp:gp|L07558|YSANAG1A_1 *Candida albicans* glucosamine-6-phosphate deaminase (NAG1) mRNA, complete cds. NID: g170885.

5 gtgtttgaagatcgacatctaactgtatggagaattaagtaaagaatttatcaggctagt
atgcgctataaaagaagtaaaattaaacaaaaagtaggtctaattggatgaacatcctgta
aataatattattaactattttgcggtacatcaaagaggtggaattccttgcatTTTTAAT
catcaatggagtaaatgaaaggatacatcaacttgtaaaaagttatgacatacaatggta
attaaagataatcatcttacctcaaatcatgataactcaatttataatgatgagggttatc
ccacgtaatgttatacatataggtttcacgtcaggaactacaggtttacccaaagcgttt
10 tatagaaatgaacattccttgatagtttcttttaaggaaaatgagaaattactccagcat
tgtgaagaaccattgtagcaccgggtcctttatcacattcactttcattgtacgcatgt
atttatgcatttaagctactggaaaaacatttataggtcaaaaaattttaatccactatct
cttatgcgtcttattaatcaattgaacaaaacgacagcaatattttagtgccaacgatg
gtacaacaacttatttcaactcaacgacattgttcacgattaaaagtattttagtagt
15 ggtgctaaacttacattgcaacagtttcaacaaatcagaaatttatatccacaagcaaat
tttaagaatttttgggacatctgaagcaagttttataagctacaattttaaccatca
tctcctgctaattctgttggttaaacttttccctcatgtcgagacacgattattaaatcaa
gatgatgatgcagtaggattattagccggttagaagtgaatgggtgttagtggttatgtt
ggacaaagcaatcaagagggggcattgattaaaacagggcgacttcgcttatattaaaaat
20 caacatttgttttagtaggttagagagagtgatcgtattatagttggggggattaatgta
tatccaacagctattgaaagcttaattatggatattgaaggcattgatgaggcccttgct
attggtataccacatgctaaatttgagaaatagcgatattgctttatcaggtaaagta
caattgaattaccgacaaattaaatcttttttaatgaaacatctttcaagacaagaagtt
ccatcaaaatttaagaaaattgaccatattgatttatacagaatcaggaaagattgctaga
25 aaagagatgaaaaataaatttattaatggagagttataa

Sequence 2942
VFEDRHLTYGELSKEIYQASMRYKEVKLNKKVGLMDEHPVNNIINYFAVHQRGGIPCI
FNHQSNERIHQLVKSIDIQWLIKDNHLTSNHDNSIYNDEVIPRNVIHIGFTSGTTGLPKAF
30 YRNEHSWIVSFKENELQLHCEETIVAPGPLSHSLYACIYALSTGKTFIGQKNFNP
LMLRLINQLNKTTAIFVVPTMVQQLISTQRHCSSIKSILSSGAKLTLQQFQQIRNLYPQAN
LIEFFGTSEASFISYNFNQSSPANSVGKLFPHVETRLNQQDDAVGLLAVRSEMFVSGYV
GQSNQEGAWIKTGDFAYIKNQHLFLVGRESDRIVGGINVYPTAIESLIMDIEGIDEALV
IGIPHAKEGFEIAILLYSGKVQLNRYQIKSFLMKHLSRQEVPSKLKKIDHMIYTESGKIAR
35 KEMKNKFINGEL*

Sequence 2943
Contig_0523_pos_0_303
is similar to (with p-value 1.0e-31)
40 >sp:sp|P14638|TRPB_METVO TRYPTOPHAN SYNTHASE BETA CHAIN (EC
4.2.1.20). >gp:gp|M35130|MVOTRPA_2 *M. voltae* tryptophan syn
thase operon (trp) genes, complete cds. NID: g150070.
atgaaaattcaaacagaagtagatgaattgggcttttctcggtgaatatgggtggccaatat
gtacctgaaacattgatgccagctattattgaacttaaaaaagcatatgaggacgcgaaa
45 tcagatactcacttcaagaaagaatttattattttaaagtgaatatgttggttagagaa
acgcctttaacatttgctgaatcatcacaaaattgttaggtgggtgccaaaatatatctt
aaaagagaagacttaaatcacactgggtgctcataaaattaataacgcgataggacaggcA
TAA

50 Sequence 2944
MKIQTEVDELGFGEYGGQYVPETLMPAIIELKKAYEDAKSDTHFKKEFNYYLSEYVGRE
TPLTFAESYTKLLGGAKIYLRKREDLNHTGAHKINNAIGQA*

Sequence 2945
55 Contig_0527_pos_178_1065
>sp:sp|P22983|PODK_CLOSY PYRUVATE, PHOSPHATE DIKINASE (EC 2.
7.9.1) (PYRUVATE, ORTHOPHOSPHATE DIKINASE).
atgggagcaacagcatttgtctataacggctcgtttccaccctgaaacatatctcgagtta
cttcaaaattatcaaatattgttctatgttgtacaccaacagaatatcgtatgatggct

aaacttagtcatttagaacagtacaatttagagtatttacacagtcggtgtctgcgggt
 gaacctttaaatacgagaagttgttgaacaatttaaactgcattttaaataactgttcga
 gatggatatggacaaaccgaaagtacattgttgatcggtatttctaaaagatactgaacca
 cgtatgggttctatgggcaaaggtataacctggtagttttgttactgtcattgacgatgat
 5 ggtaaagaggttggtccaaatgttaaaggaatatcgccgtgccttttagacttaccggct
 ttatttaaaggttactttaagatgaagcacgcacaaaagcagcttcaacaggtgattat
 tatgttactggagaccaagctcatattgataatgatggttatttctggttcgaaggcgc
 cgtgacgatattatcattagttcaggtatataccattggaccttcgaggtagaagatgca
 ctaacaaatcacgcagctgttaaagaatgtgcagttggtgcaagtcctcatgacattcgt
 10 ggaaatattgttaaagcatttatcatcttgcaagatgattatgaagcaagtgatgagtta
 atccaagaattacaagtatttgtaaaaatgaagtagcaccgtataaatatccaagagca
 attgaatttgttgaaactctacaaaaacaaattcaggtaaagatcgtcgtgttgattta
 cgtgacgcagaaataaaaaataataacaacaagattcatcacattaa

15 Sequence 2946

MGATAFVYNFRFHPETYLELLQNYQINVLCCTPTEYRMMAKLSHLEQYNLEYLHSAVSAG
 EPLNREVEQFKRHFNITVRDGYGQTESTLLIGFLKDEPRMGSMSGKIPGSFVTVIDDD
 GKEVGPNVKGNIAPVLDLPALFKGYFKDEARTKAASGDYVYVTDQAHIDNDGYFWFEGR
 RDDIIISSGYTIGPFEVEDALTNHAAVKECAVVASPHDIRGNIVKAFIILQDDYEASDEL
 20 IQELQVFCNEVAPYKYPRAEFVEHLPKTNSGKIRVELRDAEIKKYKQDSSH*

Sequence 2947

Contig_0527_pos_1455_0

is similar to (with p-value 3.0e-67)

25 >gp:gp|AF068246|AF068246_1 Mus musculus SA protein mRNA, co
 mplete cds. NID: g3928675.

atgaaagacttacttggtggttaaaggtgccaatctttcagagatgaagagactcggacta
 ccagtaacagatggtttttacaattacgactgaagcttgattacatattttaaactaaat
 gaagaactacctacagaagtaaaagacacaattaattgatcatttagcagctttttctaaa
 30 cgaacaggaaaagccttttctctgatgataacttggtatttagtatcagtagcgttaggt
 gcaaaaatttccatgccaggaatgatggatacaattcttaatttaggacttaatgatgat
 aatgtaaaaaagcttgggtgacaaaacaaatgatgcagatttgcatatgattggtaccgt
 cgtttactacaaatggttgagaggttggttatggtattccaatgacagctttcgataca
 tattttatgatgttttaaaacaaagcatcgttatcaaaatgatgctgaaattccggcagaa
 35 ggactacaaactatatgtgaaaaatataaagaaatctatgtagaagaggcatataaacct
 tttccccaagaaccgttaaagcaattagaagaagcaattgaagcagattttaaactctgg
 gataatgatcgtgcacgtgtatatagagatttaaatgatattccacatgatattggtaca
 gccgtaaatattcaggaaatggtatttggttaaatagtggtgaaaatagtggtacgggtgtg
 gcatttacaagaataccagttactggagaaaatcatttattcggagagtagtacttaatt
 40 gctcaaggtgaagatggttggtgcaggcattcgtacccttaaggatattgacactttaaaa
 caacaaatgccagatgtacatcaagagtttggtgatgtaaccaaacacttgaaaaacat
 tacaagatagcaagacatagaatttacaattgaaaatggttaaactttatttattacaa
 acacgtaacggaaaacgtactgctaaagctgcaataaaaattgctgtggatttagttcac
 gagcaattaattacacgtgaagaagcagtatcaaaagtagaggtaaaatcaatagaccaa
 45 ttattacatcctaatttttaatagaagaatcattaaagcaagcgacagtggtttctaaaatg
 ggcttaccagctagtcctcggtgcagcaacaggaaaagtagtcttctctgctgaagaagca
 aaacttcaagctgaaaatggttaataaagtagtggttaatgcgaccggaaacatcacctgaa
 gatattgaggggatggttagcaaggaatgggcaaatgttggtgacaggatggttcgaatgta
 50 gagatagatcacagtgaacaaaacagtatattatcctgaaggtgaattacatgaaggggat
 atcgtttctgtgatggttcagctggtgatttatatttaggagcaattgaaacagtcatt
 gctgaacatagtgaaagagttcgatcaatttatgacttggtctgaagagattgcaagactg
 caagttagaatgaatgctgaaacaccacaagatataaaagctggatataattttggttct
 aaaggaataggttttagttcgtactgagcacatgttcttggccctgaacgtttaatagaa
 55 atgcgccgttttatcttagcttcaaactcatgacgaacgtgtacaagctttagaaaaaatt
 aaacataccaagtagaagattttgaaacaattttcagattatctcaagatagacctacg
 attgttcgtttacttgatccaccgttacatgagttcttaccatcatctgaagaagatata
 aacaatgtttctcaacagctgaatgtatcttcagagttcttaacgcaagcgaatcgttgac
 ttaaatgaggtcaattcaatgcttggtcatcgtggttgctgtgacttatcca

gagttatatgagatgcaagttgaagctatcattgaaagtgttattaagcttcaaaaagag
 ggcataacgtgcctaccagaaattatgattcctctcgtgtcaacagtagaagaatttaca
 actttaaaagaacgatttagttaataacaattacacatttagaaaaagaatcacaacaagat
 atacaatatatgataggtactatgattgaaacgcctagagcatgcttgattgccaatgac
 5 cttgccaacattgtgatttcttcagttttggtactaatgatttaacgcaattgacattt
 ggtttctctagagatgatgcaggaaaattcataaattgtgtatactgaaaataacatttta
 cagcttgaccattccaaacttttagatagagaaggtgtaggacgactaattcaattagct
 gttgaacaagctaaaaatacaaatccagagataaaaaattggtgtattta

10 Sequence 2948

MKDLLGGKGANLSEMKRLGLPVPDGFITTEACITYLKQNEELPTEVKTQLIDHLAAFSK
 RTGKAFSSDDNLLLVSVRSGAKISMPGMDTILNLGLNDDNVKKLVDKTNDARFAYDCYR
 RLLQMFGEVYVYIPMTAFDITYFNDFKTKHRYQNDAEIPAEGLOTICEKYKEIYVEEAYKP
 FPQEPLKQLEEAIEAVFKSWDNDRARVYRDLNDIPHDIGTAVNIQEMVFGNSGENSGTGV
 15 AFTRNPVTGENHLFGEYLLNAQGEDVVAGIRTPKDIDTLKQQMPDVHQEFVDVTKQLEKH
 YKDMQDIEFTIENGKLYLLQTRNGKRTAKAAIKIAVDLVHEQLITREEAVSKVEVKSIDQ
 LLHPNFNEESLKQATVVSKMGLPASPGAATGKVVFSAEAKLQAENGKNKVVLMRPETSPE
 DIEGMVASEAIVTTHGGMTSHAADVARGMGKCCVTGCSNVEIDTVNKTVYYPEGELHEGD
 IVSVNDSAGDLYLGAIETVNAEHSEEDQFMTWSEEIARLQVRMNAETPQDIKAGYNFGS
 20 KGIGQVTEHMFPGPERLIEMRRFILASNHDervQALEKIKTYQVEDFETIFRLS DRPT
 IVRLDPLPHEFLPSSEEDINNVSQQLNVSSEFLRKRIVDLNEVNSMLGHRGCR LAVTYP
 ELYEMQVEAIIESVIKQKEGITCLPEIMIPLVSTVEEFTTLKERLVNTITHLEKESQQD
 IQYMIGTMIETPRACLIANDLAKHCDFFSFGTNDLTQLTFGFSRDDAGKFINVYTENNIL
 QLDPFQTLDRGVRLIQLAVEQAKNTNPEIKIGVFX

25

Sequence 2949

Contig_0528_pos_3612_3271

No hits found

atgtcaaatagcgcttcaatatctcagagcaacgtcgcaagtcgaagtactacagcaagt
 30 ttgagccaatctgaatcagcaaatgattcaatgagttcatctctgtccgagtcctaactca
 ataacatccgaaagtaatacaaatagcaaatcggaattgaatcaaaaagtagctctaca
 agcgagttctgtcagaatcaggaagtgtatctaactcagaaaaatctgagtcgaatttct
 cattctcaatcaacatcagctacaccttcttctcaatcgacttaccacaacaacctaata
 gaagagaagaagggttcttctgcacgtctatttaacttataa

35

Sequence 2950

MSNSASISQSNVASQSTTASLSQSESANDSMSSSLSESNSITSESNTNSKSEIESKSTST
 SEFLSESGSVSNSEKSESISHSQSTSATPSSQSTYQQQPKEEKGFFARLFNL*

40 Sequence 2951

Contig_0530_pos_2171_1191

is similar to (with p-value 3.0e-28)

>gp:gp|Y14325|ATY14325_1 Arabidopsis thaliana mRNA for meva
 lonate diphosphate decarboxylase. NID: g2288886. >gp:gp|AC00
 45 5499|ATAC005499_10 Arabidopsis thaliana chromosome II BAC T6
 A23 genomic sequence, complete sequence. NID: g3785992. >gp:
 gp|Y17593|ATH17593_1 Arabidopsis thaliana MVD1 gene, exons 1
 to 9. NID: g3250735.

gtgaaaagtggcaaacgacgagcacatacaaatattgcgttgattaagattgggggaaa
 50 gctgatgaaacttacattatctctatgaataatagtttatcagttaccttagatagattt
 tatactgaaacaaaagtgcatttgaccctgattttactgaagattgccttattttaaat
 ggtaatgaagtgaatgccaaagagaaagaaaagattcaaaactatatgaatatagtgaga
 gatttggctggaaatcggttgcagtcgcgaattgaaagtgaatttatgtgccaacagca
 gcaggacttgcttcttcagcgagtgcttacgctgcttttagctgcgcttgtaatgaagct
 55 ttgtcattgaacttatcagatacagacttatcacgattagctcgacgtggttcaggttct
 gcttctagaagatttttgggtgatttgcggaatgggaaaaagggcatgatgatttaact
 tcatatgcacatggtattaattccaatggttgggaaaaagatttatcaatgatatttgta
 gtgattaacaatcagtcaaaaaaagtatctagtaggtcaggaatgtcactaacaagat
 acutctcgatttttatcaatattggttggtcacgcttgatgaagatttaaatgaag.aaaa

gaggcagtcacaaatcaagattttcaacgcttaggagaagtcattgaagcaaatgggtta
 cgtatgcataccacttaggcgctcaacctcctttcacgtatttagtgcaagaaagc
 tacgatgctatggcgattgtggaacagtgctgaaaagccaatttaccttggtactttaca
 atggacgcgggtcccaatgtaaaagtttttagtagaaaagaaaaataacaagctgtgatg
 5 gaacaatttttaaaagattttgacgaatcgaagattatagcaagtgatatcattagctct
 ggtgttgaaattattaagtaa

Sequence 2952

VKSGKARAHTNIALIKYWGADETYIIPMNSLSVTLDRFYTETKVTFDPDFTEDCLILN
 10 GNEVNAKEKEKIQNYMNIVRDLAGNRLHARIESENYVPTAAGLASSASAYAAALAAACNEA
 LSLNLSDTDLRLARRSGSASRSIFGGFAWEKGGHDDLTSYAHGINSNGWEKDLMSIFV
 VINNQSKVSSRSGMSLTRDTSRIFYQYWLHDHVEDLNEAKEAVKNQDFQRLGEVIEANGL
 RMHATNLGAQPPFTYLVQESYDAMAIVEQCRKANLPCYFTMDAGPNVKVLVEKKNQAVM
 EQFLKVFDESKIIASDIISSGVEIIK*

Sequence 2953

Contig_0532_pos_1713_3254

>sp:sp|P54715|PTIB_BACSU PTS SYSTEM, ARBUTIN-LIKE IIBC COMP
 ONENT (PHOSPHOTRANSFERASE ENZYME II, BC COMPONENT) (EC 2.7.1
 20 .69). >gp:gp|Z99108|BSUB0005_89 Bacillus subtilis complete g
 enome (section 5 of 21): from 802821 to 1011250. NID: g26330
 55. >gp:gp|D50543|D50543_3 Bacillus subtilis DNA for 76-degr
 ee region, complete cds. NID: gl486240.

atgtttgcttttttcggtattgttttgggattcgctacattatttaaaaatccaaccatt
 25 atgggaggattagctgatcagcaaacattttggtttaaattttggtctgttattgaaatca
 ggtggttgggttaatatattacacatatggaaattgtctttgtagttggcttaccattatct
 cttgctaaaaaagcaccaggacatgcagcttttagcagctctaattgggataatttaattgttt
 aatacttttatcaatgcaattttaactcaatggccacatacttttggcgctaatttaaaa
 aaagggtgtagaaaacacaacaggattaaaatcgattgcaggtattgaaacggttagatacc
 30 aatatttttaggtgcaatcattatctcaggaataataacgtggatacataatagata:tac
 agtaagcggtttacctgaaatgttaggtgtatttcaaggattaacattcggttgtaacaatc
 tctttctttgtcatgtcaccagtgccagcaatcacttgtgtgttttgcccaacgattcaa
 cacggtattgttcaattcaatattttattgttgcacaggttatataggtgtttggtta
 tatcattttcttagagcgtgtactaatacctacaggattgcacattttatctatgcacca
 35 atcgaagtaggtccagtagttgttaatcatggtttgaaagcagaatggcttcaacactta
 aaccagtttgccgaaagtaataaaccacttaagaacaattcccatatggatttatgttg
 caaggaaatggaaagggtttttggtgtctaggtatagcattggcaatgtatgcaactaca
 ccaaaagaaaatcgtaaaaaagttgctgcattattaataccagcaacacttacggcagta
 gtacgtggttattacagaaccacttgaaattacattcttattttatttgcgccatttttattc
 40 gtattacatgcactactagcagcaactatggatacactgatgtatggatttgggtgtgta
 ggtaatatgggtggcggtgtactagattttattgcaactaactggataccattaggaaaa
 gcacattggatgacatatgtctttcaagtagtaattggttttaattctttgttgcaatttac
 tacttcttattttaaatatttaattttaaaatttgatattccattaccaggacgtaagaaa
 ggtgaagaagaagttaaattattttccaaacaagattataaagataaaaaaggagattca
 45 actcgcaatcattcacctaataagtgaaatgaagaaaaggcgatgtactacctagaagggt
 cttggaggaaaagaaaatattaaagacgttacaattgtacgacacgtctacgtttaact
 gttaaagacgaaagtaagttcaagaaagtcctattttacacataatcaaattgtctcat
 ggtttagtaaaaaagtgcaaaagtggttcaagtcgtcggttggaatgtctgtacctcaggta
 agagaagcatttgaaaatattgtaaatgatgatctatcttag

Sequence 2954

MFAFFGIVLGFATLFLKNPTIMGGLADQQTFWFKFWSVIESGGWVIFTHMEIVFVVLPLS
 LAKKAPGHAALAALMGYLMFNTFINAILTQWPHTFGANLKKGVENTTGLKSIAGIETLDT
 55 NILGAIISGIIITWIHNRYYSKRLPEMLGVFQGLTFVVTISFFVMLPVAAITCVVWPTIQ
 HGIASIQYFIVASGYIGVWLYHFLERVLIPTGLHHFIYAPIEVGPVVVNHGLKAEWLQHL
 NQFAESNKPLKEQFPYGFMLQGNGKVFGLGIALAMYATTPKENRKKVAALLIPATLTAV
 VAGITEPLEFTFLFIAPFLFVLHALLAATMDTLMYGFVGVNGMGGGVLDFIATNWIPLGK
 AHWMTYVFQVVIGLIFVAIYYFLFKYLILKFDIPLGRKKGEEVVKLFSKQDYKDKKGS
 TRNHSNPSEYEEKAMYYLEGLGGKENIKDVTNCTTRLRLTVKDESKVQESAYFTHNQMSH

GLVKSGKSVQVVVGMSVPQVREAFENIVNDDL*

Sequence 2955

Contig_0533_pos_2932_4371

5 is similar to (with p-value 7.0e-20)

>gp:gp|AB005556|AB005556_1 *Plectonema boryanum* DNA for NADP H:protochlorophyllide oxidoreductase, hypothetical protein, partial and complete cds. NID: g3123724.

atggaggtgtttaacatgacaaatcaattatttattaacaatgaatttatagaaagtcag
10 tctaaagagacaatggatgtcattaatccagctactggcgaggcatttgatactatcact
cttgcaactgaagaggaagtaaacgacgccattgaaaaatcgcaacaagcacaacttgaa
tgggagcgcgtgcctcaacctacacgtgcggaacatgttaaattactttatcctttatta
gaaaaaaatcgcgatgaaatagctcaatttatcgtaaaaagaacaaggtaaaacttttagcg
caagcttatggagaaattgacaaatcaatctcatttatcgatttatgacaagtctgagt
15 atgtcagataaaggacgtgttctacaaaatagtattgcaaatgaaacgattcaaattatc
aacaacctatcggagttactgctggtattgtgccatggaacgcaccgatacttgcctt
atgcgaaaagtcattccagctatagtaactggttgcagtagtgattaaacctagtga
gagacaacgttactcactcttcgattagctgaattatcgagagcatcaactataccagca
ggttggtttcaaatgttctcctggcactggagaacagtaggtacacaattagcttcgcat
20 aaagacattcaacttatttctttaactggaagtatgagagctggtaaatctgtttacgaa
aatgctgctcaaactgtaaagaaagtaaattagaattaggtggaacgcaccagtcatt
gtcacatcaaagtcgatttagataaagcagttgactatcgtgacagcccgataaat
aatgcaggtcaagtttgtagctgccctgaacgtatcttctgacatgaagatgttcacgat
gactttttaataaaagtaacttccaaaatgaaaagcttaactggtggagatccatttgat
25 gaaaacaccgattacggcgcaattattaacaaaaacaacttgatagattcatgaaaag
gttcaagatgctattaaaaatggtgcaacattgatgactggtggacatcaattaaaacgc
catggtttcttctacgcaccaacagttatagataacgttagaaaagacgataatgtgtt
aaagatgaaattttgggtccagttcttgcgataacaacctaccgtgatattgaacaagt
attgaagacgctaataatgatacaaacgctggcttatcttcttatcttctctgaaaattta
30 acagaagtatgacagcaaccgaacgtctaaaatttgggtgaagtatatgcaaatgtgag
gctgaagaagtggtaaatggctatcacgcaggttggcgatgaatcaggcttaggtggcgct
gacggtattcacggttttgaagaatactacaataaccacagtaagttatattagatactaa

Sequence 2956

35 MEVFNMTNQLFINNEFIESQSKETMDVINPATGEAFDTITLATEEEVNDIEKSQQAQLE
WERVPQPTRAEHVKLLIPLLEKNRDEIAQLYVKEQGKTLAQAYGEIDKISFIDYMTSL
MSDKGRVLQNSIANETIQIINKPIGVTAGIVPWNAPILVLMRKVIPAIVTGCSVVIKPS
ETLLTLRLAELFRASTIPAGLFQIVPGTGETVGTQLASHKDIQLISLTGSMRAGKSVYE
NAAQTIVKKNLELGGNAPVIVTSNADLDKAVDYIVTARINNAGQVCTCERIFVHEDVHD
40 DFLNKVTSKMKSLTVGDPFDENTDYGAIIINQKQLDSIHEKVQDAIKNGATLMTGGHQLKR
HGFFYAPTFLDNVRKDDNVFKDEIFGPVLAITTYRDIEQVIEDANDTNAGLSSYIFSEN
TEVMTATERLKFGEVYANCEAEVVNGYHAGWRESGLGGADGIHGFEYYNTTVSYIRY*

45 Sequence 2957

Contig_0533_pos_8501_7722

is similar to (with p-value 2.0e-98)

>sp:sp|P25553|ALDA_ECOLI ALDEHYDE DEHYDROGENASE A (EC 1.2.1.22) (LACTALDEHYDE DEHYDROGENASE).

50 gtgggttacgggtgtagcacaagcaggtgctgaacgtatcggttcctacaacattagaacta
ggtggttaaaagtgtcaatattatctttgatgatgctaatttagagcaagtgaaggt
gttcaacttggtatattatttaaccaaggtgaagtatgtagtgcaggttcaagattatta
gtacaatcatctatttacgatgaattgttgccaaaattgaaagaagcatttgaaaatatt
aaagttggcgatccatttgatgaagataactaaaatgagtgcgcaaaactggaccagaacaa
55 ttagataaaaattgaaagttatataaaaattgctgaagaagtacaaaagcgaacatctta
actggtggtcatcgaatcacagataacggcttagacaaaggttacttctttgagcctaca
attattgagattaacgataacaagcatcaacttgctcaagaagaaatcttcggtccagtt
gtagtagttgaaaaattcgatgatgagcaagaagctatcgaaattgcaaatgattctgag
tatggttttagctggaggtatcttctactacagatattcatcgtgcattaaatgtagctaaa

gctatgagaacaggtcgtatttggattaatacttataatcaaattcctgctggtgcgcc
 ttcggaggatataaaaaatcaggtattggcgcggaagtatataaagatgctatcaaaaac
 tatcaacaagttaaaaatatctttattgatacaagcaaccaaaactaaaggtttatattaa

5 Sequence 2958
 VGYGVAQAGAERIVPTTLELGGKSANII FDDANLEQVIEGVQLGILFNQGEVCSAGSRL
 VQSSIIYDELLPKLKEAFENIKVGDPFDEDTKMSAQTGPEQLDKIESYIKIAEEDDKANIL
 TGGHRITDNGLDKGYFFPTIIEINDNKHQLAQEEIFGPVVVVEKFDDEQEAIEIANDSE
 YLAGGIFTTDIHRALNVAKAMRTGRIWINTYNQIPAGAPFGGYKKSIGREYVKDAIKN
 10 YQQVKNIFIDTSNQTGLY*

Sequence 2959
 Contig_0533_pos_0_1310
 is similar to (with p-value 7.0e-62)
 15 >sp:sp|P40047|DHA3_YEAST ALDEHYDE DEHYDROGENASE, MITOCHONDR
 IAL 3 PRECURSOR (EC 1.2.1.3). >gp:gp|U56605|SCU56605_1 Sacch
 aromyces cerevisiae precursor aldehyde dehydrogenase gene, n
 uclear gene encoding mitochondrial protein, complete cds. NI
 D: g1336077.
 20 gtgtttaatatgtcattggaattgccggtgatgctaacgattgtgttgttcttagcacta
 ggcatttttagtcaatggtagcgagtagaataaaatggccatcgattgttgcattggcc
 atcgtaggtttactttaggacctatttttggattagcaaatccaaaagaggcacttgga
 cctgaggcatttagttcaattgtatctcttgcgtgtagcaattatattattgaaggtagt
 agtaacttagatttttagagaattaaaaggcatttctaaagctgttataagaattattaca
 25 ataggagcgggaattgcatggatttttaggagcaatcgctttacatgtcactatgaatttc
 cctctgtctatttcatgttgcggagggtattcctaatactggtccaaccgtgatt
 caacccttgcgaagcaagcgaaagtaaaaagaaatgtcgattcagtattaagatgggaa
 agtattatcttagatcctattggacctattattgcactcactgcattttatgtattccag
 atttttgaggaagggataggctttgtcgttattattctctttatcttgaagctcttagct
 30 gcaattttaatagggttttggcgagcctttctatttaattggcttataggccaagataaa
 attcctcaaagttaatgccaccattcagtttgtatttattcttttaacatttagtattc
 tgcgtagatattatcagaatcaggtttactagctgtaacaatttttgggttaattgatg
 gcgcgtaaaaaacgtcacgatctcatctttaaagaatcagaccatttatagataa:gca
 tcatccatcactgtgagtactgtatttattttaattacgtcttctctcactaaagtgtg
 35 ttgctcaatgtgttatcttggcagctcactctttagtttgggttatgattgtattagta
 agaccaatttcagtgtttcttcaacattaggtactgaaataactaaaaaggaacgtgca
 gtagtagcactaatggcgctagaggtattgtggttttaacagtggcacaattcttctca
 agtttatttatggacgataaaaattcctatggctcaatatattacaccagtaacttttgg
 cttgtatttatcactgtagtcatttatggatttgggttttacaccttaagtaaaactgttc
 40 ggtgtagcaagtacggagccaccaggcgtaatcattgtcggagaaagtgaattttcgttc
 catcttgggtattaatctaagggtatcatggtataccgctcatgatgttcaa

Sequence 2960
 VFNMSLELPVMLTIVLFLALGIFSQWLASRIKWPSIVVMAIVGLLVGPIFGLANPKEALG
 45 PEAFFSSIVSLAVAILFEGSSNLDRELKGIKAVIRIITIGAGIAWILGAIALHVTMNF
 PLSISFVIGLFLITGPTVIQPLLKQAKVKNVDSVLRWESIILDPIGPIIALTAFYVFQ
 IFEEGIGFVVIILFILKLLAAILIGFGAAFLFNWLIGQDKIPQSLMPPIQFVFILLTFES
 CDEILSESGLLAVTIFGLMMARKKRHDLIKESDHFIDNASSILVSTVFILITSSLTQDV
 LLNVLSWQLILFSLMIVLVRPISVFLSTLGTEITKKERAVVALMAPRGIVVLTVAQFFS
 50 SLFMDDKIPMAQYITPVTFLVFITVVIYGFQFTPLSKLFGVASTEPPGVIIIVGESEFSF
 HLGINLRDHGIPVMMFX

Sequence 2961
 Contig_0534_pos_7490_6273
 55 is similar to (with p-value 3.0e-99)
 >sp:sp|P39312|CYCA_ECOLI D-SERINE/D-ALANINE/GLYCINE TRANspo
 RTER. >pir:pir|S56433|S56433 hypothetical protein o470 - Esc
 herichia coli >gp:gp|U14003|ECOUW93_120 Escherichia coli K-1
 2 chromosomal region from 92.8 to 00.1 minutes. NID: g126317

2. >gp:gp|AE000492|AE000492_4 Escherichia coli K-12 MG1655 section 382 of 400 of the complete genome. NID: g1790649.
atgttcacatgagagctatgggagaattactgttatccaatttaggatttaaactcggttggt
gacattgctcatcatcatattgggttctatggcaggttttatgggtgggtggacatattgg
5 ttaacatggattatttcaggaatggcagaagtgactgctggtgccaagtatggttcccttc
tggtatccaacaattccaaactggtaacagctgcagcgactatttttagtttagttgct
ttaaatctattcagtgctaaattatttggagaattagaattttggctatctattattaaa
gttttgactatttttagctttgatagccgttggtgtgttatgattgtatttggaaatgaag
acaagctatggccctgcaacggttaacgaatatatggaaagacggaggcttttccctaatt
10 ggtgcacaaggtttcttcattgtcattccaaatggcaattttctcattttattggtattgag
ttgattggaataactgcaggggagactaaagatcctcacaacaattcctcaagcaatt
aataatgaccgttttagaataattattttatataaggatcggttggaagttatcatgtct
gttgatccatggcaacaattgaatcctgctgacagtcacatacgttaaaatgtttggatta
gttggaaatcccttttgcagcaggtattattaactttgttgacttacagctgcagcctct
15 tcttgtaatagtggtatatttgcataatagccgtacgatgtttggattagctggaagaaag
caaggtccagcattcttacatagaaccaataagcacggcgatccacattatgctatttta
gtgacatgtggttatttaagtatttcagtcgtgttaaatgcaatttttaagatgcgact
aaagtgttcgtacaataattacaacattttcaactgttttaaatattatgatttggacaatt
attatgctgcgctatctaggttatttaagacatgaaccgaaacagcataaagaaagtaac
20 tataaaatgtggggcggaataacatggcttacagtttttaggggtctttgcatttatt
tttattatactattgattaatagtgaacgcgttatgccgtactttctgcacccgtatgg
tttgattatcatgctattgatgtatcaaaaatataaaaaagaatctcgcaagctaaaatt
aaaaatgaggaagagtaa

25 Sequence 2962
MEMRAMCELLSNLGFKSGFDIAHHHIGSMAGFMVGWYWLTIISGMAEVTAVAKYVSF
WYPTIPNWLTAATILVLVALNLFSAKLFGLEFWLSIIKVLTI LALIAVGVMIVFGMK
TSYGPATVTNIWKDGGFFPNGAQGFMSFQMAIFSGIELIGITAGETKDPHKTI PQAI
NNVPRILLFYIGSLAVIMSVVPWQQLNPADSPYVKMFGVLGIPFAAGIIN FVVLTAAS
30 SCNSGIFANSRTMFLAGRKQGPFLHRTNKHGVPYAILVTCGLLSISVVLNAIFKDAT
KV FVQITTFSTVLNIMIWTIIMIAYLGYL RHEPKQHKESNYKMWGCKYMAYSILGFFAFI
FIILLINSATRYAVLSAPVWFVIMLLMYQYKKESRKAKIKNEEE*

Sequence 2963
35 Contig_0539_pos_0_15071
is similar to (with p-value 8.0e-35)
>gp:gp|X81475|MHLMP_1 M.hominis lmp1 and lmp2 genes. NID: g
587470.

atgggcacgttaaaatcattagttgctaaacaacctacagtacaaaaacaagtgtttat
40 attaacgaagatcaacctgagcaatctgcctacaatgattccattacaatgggacaaact
ataattaataaaacagctgatccagtacttgataaaacttttagttgataacgcaatcagt
aacatttcaactaaagagaatgcactgcatggtgaacaaaaattaacaactgctaaaacg
gaagcaattaatgcacttaatacattagctgatttaaacacacctcagaaagaggctatt
aaaacagctattaacactgctcatacaagaactgatgtaactgcagagcaaaagtaaggct
aatcaaaataatagtgcaatgcacacgttgagacaaaacatttctgacaacgaatcagta
45 acaaacgaaaagtaattatattaacgctgaacccgaaaaacaacatgcctttactgaggct
ctaaataatgctaaagaaatagttaatgaacaacaagccactcttgatgccaatcaatt
aaccaaaaagcacaagcgattcttactactaaaaatgcttttagatggtgaagaacaatta
cgctcgctgctaaagaaaatgccgatcaagaaatcaatacgttaaatcaattgactgatgcg
50 caaagaaatagtgaaaaagggttagtcaacagttctcaaaactagaacagaagttgcttct
caattagcaaaaagctaaagaactaaataagggtgatgaacaactgaataaccttatcaat
ggtaaaaaccaaattgataaatagcagtaatttatcaatgaagatgcgaaccaacaacaa
gcatattcaaatgcatgcaagtgacagagtgcttaaaaaacaattcacaacaaacctgaa
ttagataaagtaacaattgaacaagcaattaataatattaattctgcaattaataatcta
55 aacggtgaagctaaactgactaaagctaaagaagatgctgttgcttcaataaacaaccta
agcggattaaactaacgagcaaaaaacaaaagaaatcaagccgttaattggctctcaaaact
agagaccaagttgctaacgttttacgtgattcaaggcatttagatcaatctatgcaaacaa
ttacgtgatttagttaacaatcaaaatgtaatacattcaacaagtaattattttaacgag
gattcaactcaaaagaatactttatgataatgcaattgataatggctcgacatatataact

ggtcaacacaattcagaattaaataaatctactattgatcaaacgattagccaaattaac
 acagctaaaaatgatttacatggtgcagaaaagttacaaagagataaggggaactgcta
 caagaaattggacaattaggttatTTAAATGACCCTCAAAAATCTGCTGAGGAATCCTTA
 gtcaacgggttcaaatacacgttctgaagtagaagagcatcttaatgaagctaaatcatta
 5 aataatgcaatgaaacaattaaagagataaagtagctgaaaagactaatgtcaaaacaaagt
 agcgattacattaatgattcaactgaacatcaacgtgggtatgatcaagcacttcaagaa
 gcagaaaatattattaatgaaatcggtaatccaacattaaataaatcggaattgaacaa
 aagttacaacaattgactgacgctcaaaatgcgttacaaaggttcacatctattagaagaa
 gctaaaaataatgcgattactgaaatcaataaacttacagcattaaatgatgcacaacgt
 10 caaaaagcaattgaaaatggtcaagcacagcagacaatcccagcagttaatcaacaatta
 actttggatagagaaataaatactgcaatgcaagctttacgagataaagtaggccaacaa
 aataaagttaccacaaagtaattatttcaatgaagatgaacaaccacaaacataactat
 gataattctgtacaaagcgggtcaaaactattattgataaaacttcaagatccaatcatgaac
 aaaaatgaaattgagcaggctattaatcaaatcaatacgactcaaacagcgttaagtggg
 15 gaaaataaattacacactgaccaagaaagcacaataagacaaatagaaggtttatctagt
 ttgaacacagctcaaatcaacgcggaagaaagatttaagtcaatcaagctaaaacaagaa
 gaigtgtctcaaaaggttagctacagctaaagaaataaattctgctatgagtaatt::aaga
 gatggcattcaaaataaagaggacatcaaacgtagcagtgcatatatcaacgcagatccg
 actaaagttacagctttacgatcaagcactacagaacgcagaaaatatcatcaatgccaca
 20 ccaaacgtagagcttaataaagctacaattgaacaacgcgtatcacgcgttcaacaagca
 caacaagatcttgatgggtgttcaacaattagctaatgctaaacaacaagctacacaaact
 gtcaatgggttaaatagcttaaatgacgggtcaaaagcgtgaattaaatctattaat
 tcagctaataccggtacaaaagtacaagaagaattaaacaaagcaactgaatcgaacat
 gcgatggaagctttaagaaacagtggttcaaaacgttgatcaagtaaaacaaagtaacaat
 25 tatgtcaatgaagatcaacctgaacagcacaaattatgataatgctgtcaatgaagctcaa
 gtacaatcaacaacaatgctcaacctgttctagacaaattagctatagaacgtttaact
 caaactgttaacactacaaaagatgcattacatggtactcaaaaactgatacaagaccaa
 caagctgctgaaactggaatacgtgggttaacgagctctcaatgaacctcagaaaaatgct
 gaagtagctaaagtaactgcagcaacaacacgtgatgaagtgaagaaatattcgtcaagaa
 30 gcaacaacattagatactgcaatgcttggtttacgtaaaagcattaaagataaaaacgat
 actaaaaatagtagtaaatatattaatgaggatcatgaccaacaacaagcttatgacaat
 gctgtaaataatgctcaacacgttatcgatgaaactcaagcaacgttaagctcagataca
 atcaatcaattggcaaatgccgttaactcaagctaaatctaatcttcatggagatactaaa
 ctacaacacgataaagatagtgctaaacaaacgattgctcaattacagaatttgaattca
 35 gctcaaaaacatatggaagattctttaattgataatgaatctacacgtacgcaagtccaa
 caagatttaacagaagctcaagcttttagatgggttaatgggtgccttaaaagaaagtatt
 aaagataataactaatattggttcaaacggtaattacatcaatgcggaacatctaagaaa
 caagcatatgatgcagctgtacaaaatgctcaaaaataataaataatggaacgaatcaacca
 acaattataaaggttaattgctcactacagcaacacaaacggtgaaaaatactaagatgcc
 40 ttagacgggtgatagattagaggaagctaaaaataatgccaatcaacaatcagaat
 ctatctaatttgaacaatgccaaaaagatgcagagaaaaatctagttaatagcgcatca
 acattagaacaagttcaacaaaacttacaaacgctcaacaattagataatgctatgggt
 gagttacgacaaagatttgctaacaaagatcaagtgaagcagatagtaaatatctaact
 gaagatcctcaaatagcaaaaactatgatgatgcagttcaacgtgttgaaactattatt
 45 aacgaaactcaaaacccgtgaattacttaaaagcaaacattgaccaagcaactcaatccgtt
 caaatgcagaacaagctttacatgggtgctgaaaaattaaatcaagacaaacaaacgtct
 tcgacagaactagatggattaacagatttaacagatgcacaacgtgaaaaactcagagaa
 caaattaacacttctaattagtagagatgatattaaagcaaaaaattgagcaagcaaaagca
 ctaaatgacgcaatgaaaaaacttaagaacaagttgcgcaaaaagatgggtgttcatgct
 50 aacagtgattatacaaatgaagattctgcacaaaaagatgcgtataataatgcactaaa
 caagcggaagacattattaataacagctcaaatcctaacttaaatgcacaagacattact
 aatgctttaaataatattaacaagcacaagataaccttcattggagctcaaaaattacag
 caagacaaaaatacaactaatcaagccattggttaacttaaatcatcttaataacactcaa
 aaagatgcgcttatacaagctattaatggagctacatctagggaacaaagttgcagaaaaa
 55 cttaaaagaggccgaagcgttgatgaagctatgaacaacttgaagatcaagtgaatcaa
 gal jntccaaatttcaaatagcagccattcataaatgaagactcagacaaacaaaact
 tataatgataaaatccaagctgcaaaagaaataattaatcaaacatctaactcaaacctta
 gataaacaacaaaatttgctgatcaacttcaaaaatattaagatgcagtgaaataatttacat
 ggtgatcaaaaattagctcaatctaacaagacgctaataatcaattaaatcatttagat

gacttaaccgaagaacaaaaaacattttaaacggttaattaataatgctgatactcga
 gatgaggttaataaacaactagagattgctaacaattaaatggtgatagagtacactt
 cataaagtcataaatgataaagatcaaattcaacatttaagcaattacattaatgctgat
 aatgataaaaaacaaaattatgataatgctattaaagaagctgaggatttaattcataat
 5 catccagatacatttagatcataaaagcattacaagatttattaacaagatagaccaagcg
 cataacgaattaaatggagaatccagatttaaacaggttttagacaatgctttaaacgac
 atagatagcttaaacagctctcaatgttccacaacgtcaaactgttaaggataacatcaac
 catgtgacaaactctagaaagtttagctcaagaattgcagaaagcaaaagagcttaattgat
 gctatgaagcaatgagagatagcattatgaatcaagagcaaaattcgtaaaaataagcaat
 10 tataactaatgaagacttagctcaacaaaatgcctataatcatgcagtagataatataaat
 aacattatggtgaagacaatgcgacgatggatcctcaaataatcaacaagcaactcaa
 gatataaatacagctataaatggattaaatggagatcaaaaacttcaagatgcaagaca
 gatgctaaacaacaaattactaactttactgggttaactgaaccacaaaaacaagcattg
 gaaaacatcattaaaccaacaaacaaagtagagcaaatgttgctaaacagttaagtcatgct
 15 aaattctttaaattggaataatggaagaattaaaagtgcagtagccaaagcgctattagta
 agacaaaatagtaactatatttaagagatgtctctgaaaaagaagcatatgaacaagct
 attgcaaaaggtcaggaaataatttaattcagaaaaataatccaacaataagtagtactgat
 atcaatcgtagcattcaagaaatttaattgatgctgaacaaaattctcatggtgaaaaataa
 ttaagacagcacaggaattgcaagaatgaaatacaaaaattctagacggattaaattca
 20 gctcaaatacaaaaatttaattcaagatataggcagaacaacaaactaaacctgcagtaact
 cagaaactagaagaagcaaaagcaataaaccaagctatgcaacaacttaacaaagtata
 gccgataaggatgctactctaaattctagtaactatctcaatgaagattctgagaaaaag
 tttagcgtagcataatgctgtaagccaagctgaacaactcataaatacaacttaacgacca
 actatggatataagtaatttcaagctatttactcaaaaggtcattcaagcaaaagattca
 25 ttgcacgggtcgaaataaacttgcaaaaatcaagcagattcaaaatttaataataaatcaa
 tcaacaaatttaaatgataaaacaaaagcaagcattaaatgacttaatttaattcatgctcaa
 actaaacagcaagtggcagaaataattgcacaagctataaagttaataaacgaaatgggc
 acacaaacaaactcgtagaagaacagtcacacgttcatcaacaaagtaaatataataat
 gaagatccgcaagttcaaaaattttataatgactccattcaaaaaggctcgagaaaattta
 30 aacggcactacagatgatgttttaacaacaataaaaatagcagatgccattcaaaaacatt
 catttaactaaaaacgatttcatatgggtgatcaaaaattacaaaaagcacaacaagatgca
 accaatgatttaactattttaacaattctaaacaattctcaagacaaaagcgagcatgat
 gagattaaactctgctccttcaagaactgaagtttctaatgatttaaatcatgctaaagca
 cttaatgaagctatgctgcaacttgagaatgaagttgctcttgaaaacagtggttaaaaaa
 35 ttaagcgactttatcaatgaagatgaagcggcacaataatgaatatagtaatgcacttcaa
 aaagctaaagacattatcaacggcggttccaagtagcacttttagataaagctacaattgaa
 gatgctttattagaattgcaaaatgctagagaaagtttcatatgggtgagcaaaaacttcaa
 gaggctaaaaatcaagctattgctgaaattgataatttacaagcattaaatcctggacag
 gttcttgctgaaaaaacatttagttaaccaagcatcaaccaaaccagaagttcaagaagcc
 40 ttacaaaaagcaaaagaacttaatgaagctatgaaagcactgaaaactgaaataaaataaa
 aaagaacaaatcaaggctgatagtagatatgtaaatgctgacagtggtcttcaagcaaat
 tacaattctcggttaaatatttggttctcaaatatttgcaactaccaaccaccagagctt
 aataaagatgtaataaatagagcaactcaaacgattaaaactgctgaaaaataatttaaat
 gggcaattcaaattagcagaggctaagtgcagcggaaatcaagcatcgaacatttgcaa
 45 ggatttaacacaatcacaaaaagataaacaacatgatttaattaatcaagctcaactaaa
 caacaggtagatgatatcgtaaaataactctaaacaattagataactctatgaatcaacta
 caacaaattgttaacaatgacaatacagtaaaacaaaatagtgatttcattaatgaagat
 tccagccaacaggatgcttataatcatgcaattcaagcagcaaaagatttgataaattgct
 catcctactatcatggataaaaatcaaatagatcaagctattgaaaatatcaacaagca
 50 cttaatgatttacacggtagtaataaaactatcagaagataaaaaagaagcttcagaacaa
 ctacaaaaccttaataagcttgacgaacgggcaaaaagatacagattttaaatacatatttc
 agtcaccaacaagaagcgaagtaggagaaaaaattgcaagtgctaaacaattaaataat
 acaatgaaagcacttagagattctattgctgataataatgaaattttacaaagtagtaag
 tacttcaatgaagattctgaacaacaaaatgcttataatcaagccgtaaataaagctaaa
 55 aatataattaatgatcaaccaacaccagtaatggcaaatgatgagattcaaaagtgccta
 aatgaagttaaacaaactaaagataatttcatatgggtgatcaaaaacttgctaacgacaag
 acagatgccaagcaacattaaatgcgttaaatfacttaaatcaagcgcaaaagaggttaat
 cttgaaactaaagttcaaaaacttaattctagaccagaagtacaaaaagtagttcaatta
 gcaaatcaacttaattgatgcgatgaaaaaattagatgatgctttaactggtaatgacgca

ataaaacaaacgagtaattatattaatgaagatacttctcaacaagttaactttgatgag
 tatacagatagaggtaaaaacatagttgctgaacaaacaaatccaaatatgtctccaact
 aatattaacactattgctgataaaattactgaagctaaaaatgatttacatggcgtacaa
 aaactagaacaagctcaacaacagtcacatcaatactattaatcaaatgactgggtctaaac
 5 caagctcaaaaagaacaattaaatcaagaaattcaacaaactcaaacccgttctgaagta
 catcaagtaattataaagcacaaagctttaaatgattcaatgaatactttacgtcaaagt
 attactgatgaacatgaagttaaacaacaagtaactacatcaatgaaactgttggtaat
 caaactgcataatacaatgccgttgatcggtgtaaaacaaataatcaatcaaacatcta
 ccaactatgaatccttttagaggtggaacgtgcaacatcaaatgtaaaaactttctaagat
 10 gcacttcatgggtgaacgtgaattgaatgacaataaaaaattcaaaaaacttttgcagtcaat
 gacttagataacctcaatcaagctcaaaaagaagcatttaactcatgaattgaacaagca
 actatagtttccaagtaataatatactataacaaagcgaaagctttaataatgatatg
 aaaaaacttaagatatcgttgctcaacaagataatgtgagacaatcaacaattatata
 aacgaggatagtagacacctcaaaaatgtacaacgatataatcatgcacaatcaatc
 15 attgatcaagtagcaaacctacgatgtctcatgacgaaatagagaatgcaatcaataac
 ataaagcatgccatcaatgcactcgatggagaacataaattacaacaagcaaaaagaaat
 gcaaaacttattgattaatagtttaaacgatttaaatgcaccacaaagagatgccataaat
 agattgggttaatgaagctcaacaagagaaaaagtagctgaacaacttcaaagtgctcaa
 gctttaaatgacgctatgaagcatttaagaaacagcattcaaaatcaatcatccgttaaga
 20 caaggaagcaaatatattaatgaagtgatgctaaaaaagagcaatataatcacgaggtt
 agagaagtcgaaaatattatcaatgaacaacatccaacattggataaagaaataattaag
 caactaacggatgctgtaaatcaagcgaataatgacttaaatggcgttgaattattagat
 gctgataaagcaaaacgcacatcaatcgatacctacattgatgcacttaaatcaagcaca
 caaaacgcattaaatgaaaaaatttaataacgcagttaccagagctaaagttgcggctatt
 25 attggccaagcaaaaatactcgatcatgctatggagaatttagaagaaagtatcaagat
 aaagagcaagtcacaacagtcagtaactatattaatgaagacctgatgttcaagaaaca
 tacaataacgcgcttgatcatgtgacagaaatacttaatacaacagtaaatccaacttta
 tctattgaagatatagagcatgctatcaacgaagttatcaagcgaaaaacaactcaga
 ggtaaaacaaaactttatcaaacatcgatttagctgataaagaattaaagtaaatggat
 30 gatttaacatcacacaaagcagttcaatatctaatacaaatatatactgctaaaacgaga
 acagaagttgccaagcaattgaaaaagcaaaatcattaaatcatgcaatgaaagcactt
 aacaaaaatataaaaaatgcagataaagtggttagatagtagtcgattcattaacgaagat
 caacctgaaaaagaggcgtatcaacaagctataaatcatgttgattcaatcattcataga
 caaacaatcctgaaatggatccaacagtaataatagcataactcatgaactcgaaaca
 35 gctcaaaaataacttacatgggtgatcagaaacttgctcatgcaaaaacagatgccgcta
 gtaattaatgggtctaattcatcttaattgttgctcaacgcgaggttaatgataaatacgaat
 acaaatgctacaacacgcgaaaaagttgcaagaacttagataatgctcaagctcttgat
 aaagctatggaaacactacaacaagtagttgctcataaaaaataatattgaacgatag
 40 aattatttaaatgaagattcaaaaatatacaacaacaatcagatcgagttattgctgaaggc
 gaacaactacttaatacagacaacaaatccaacattagaacctataaagtcgatattgtt
 aaggataatgtcctagctaacgaaaaataactatttggcgcagaaaaactatcatatgac
 aatcaaatgcaaatgatgaattaaacatatagaattatcttaataatgcacaaaagcaa
 tctataaaaagatatgatttctcacgcagcattaagaactgaagttaaacaacttctgcaa
 caagctaaaacccttgatgaagctatgaaatcacttgaagataaaaactcaagtagtgatt
 45 acagatactactttgcctaattacactgaagcttcagaggataaaaaggaaaaagtagac
 caaactgtatcacatgctcaagcaatcattgataaaaataaatggctcaaatgtaagttta
 gatcaagtagcagaagcactagaacaatttaactcaagcatcagaaaacctcgatgggtgat
 cagcgagttgaagaagctaaagttcatgctaatcaacaattgaccaattacacatctt
 aattcattacaacaacaaactgcgaaagaaagtgtaaaaacgcaacaaactagaagaa
 50 atcgctactgctagtaacaatgctctggcattaaacaaagtaattgggttaattagaacaa
 ttcattaatcatgctgattctattgaaaatagtataattatagacaagccgacgcagac
 aaaattatcgcttagtgatgacatgaacatggacaagatatcaaaaaatctaacgca
 acccaaaatgaagcaaaaacagcgcttaacaacaattataaaatgcagaaacatcgtaaat
 ggtttcgaaagattaaatcatgctagaccagagctttagaatatattaaatcactagaa
 55 aaaaataacaatgctcaaaagctgcttttagaggataaagtaacgcaatcgcatgattta
 ttagaattagaacatcttgcacagagggcacaaacctcaatgacattatgggtgaatta
 gtaaacgcaatcgtaataactatgctccaacaaagcaagtaataattatattaacgcc
 gataacctacgcaaaagataactttactcaagctatcaacaatgcacgtgatgcactcaac
 aaacacgaaggtcagaacttagatttcaatgcaattgatacatttaaagatgataatttc

aaaactaaagatgcacttaacggtattgaacggttaacagctgcaaaatcaaaagcagaa
 aaactaattgatagtttaaaatttattaataaagctcaattcacacatgcaaatgatgaa
 attatgaataactaatttctattgcacaattgtctagaatcgtgaatcaagcatttgattta
 aatgatgcaatgaaatctttaagagatgaacttaataatcaagcttttctgtccaagca
 5 agctcaaattatataaattcagatgaagatttaaaacaacaatttgaccatgctttaagt
 aatgctcgaaggtacttgcaaaagaaaatggtaaaaatttagatgaaatacaaatgag
 ggactcaaaacagtgattgaggatactaaagatgctttaaatggatccaacggttatca
 aaagctaaagctaaagcaattcaatacgtacaatctttatcttatatcaatgatgcacag
 10 cgtcatattgctgaaagtaataattcacaactctgatgatttatcatcttttagcaaataca
 ttatctaaagctagtgatttagataatgcaatgaaagacttacgagatactctagaaagt
 aattcaacttctgttccaaatagtggaatttatattatgctgataagaatttacaatt
 gaatttgatgaggcgctacaacaagtgcaacaagttctaaaacttcagaaaatcca
 gcaacgattgaagaagtattaggtcttagtcaagccatttacgatacaaaaaatgcattg
 aatgggtgaacaacgctcttgcaactgagaagagcaaaagatttaaaattaataaaaggatta
 15 aaagattttaataaagcacaacttgaagatgtcacaacaaggtaaattcagcaaatact
 ttaacagagttatctcagctcactcaatcaacggttaaaatataacgataaaatgaaatta
 ttgagagataagcttaaaaccttagtaaatctgtttaaagcaagtttaaatatagaaac
 gctgattataattttaaaccgtcaatttaacaaagctttaaaagaagctaaaggcgatta
 aataaaaaatagcggtaacaattgtcaatatcaatgacattcaacatcttttaacacaaata
 20 gataatgctaaagaccaattaaatgggtgaacgacgtctaaaagaacatcaacaaaaatct
 gaagtatttattattaaagaattagatatacttaataatgctcaaaaagctgcaataatt
 aatcagattagagcgctctaaagacattaaaataattaatcaaatcgttgataatgcaata
 gaattaaatgatgctatgcaaggtttaaaagaacatgtagctcaattaacagcaactaca
 aaagacaacattgaatattttaaagctgatgaagaccttaaaatacaatatgattacgct
 25 atcaacttagcgaataatgttcttgacaaaagaaacgggtacaaataaagacgctaataatc
 ataattggaatgattcaaaacatggatgatgctagagcacttctaaatggaattgaaaga
 cttaaagatgctcaaaacaaagcacataatgacattaaagatacgtcaaacgtcaactt
 gatgaaattgaacacgctaataacacatcaaatcttaaagctcaagcaaaacaaatggta
 aatgaggaagctagaaaagcgttttctaatattaatcacgcaacatcaaatgatttagtt
 30 aatcaagcaaaagatgaagggcaatctgcaattgaacacatacatgcagatgaattacct
 aaagcaaaactagatgctaatacaaatgattgacaaaaggttgaagatataaatcaattta
 attagtcaaaatccaaacttatcaaatgaagaaaaaaataaaactaatatctcaaatat
 aagttagtaaatggaattaaagaatgaaattcaacaagctataaaacaaacaaatagaa
 aatgctacaacaaaactagatgaagtcattgaaactactaaaaaattaatatcgccaaa
 35 gcagaagctaaacaagtgataaaagagttatcacaagaagaaacgagatgcaataaataac
 aacactgatttaacaccttctcaaaaggcacatgctttagcagatattgataaaacagaa
 aaagatgcacttcaacatactgaaaatttctaattcaattgatgatatacaatacaataaa
 gagcatgcatttaatactttagctcatatcattatttgggatactgatcagcaaccatta
 gtttttgaaactacctgaattgagccttcaaaatgctctagtaacaagtgagggtggttgt
 40 cacagagatgaaactatttcattagaatctataattggagctatgactttaactgatgaa
 cttaaagtcaatattgtttcattaccgaacactgataaagtagctgatcacctaaccgct
 aaagttaaaggttatttttagctgatggctcatttgtcactgtaaatgttcagtcagagtt
 gtagaaaaagaattacaaatagctaaaaaggatgctataaaaaaattgatgttctggta
 aaacaaaaaatacaagatatagatttctaataacgaattaacgtctactcaacgtgaagat
 45 gcaaaagctgaaattgaaagattgaaaaagcaagccatcgataaagtgactcattctaaa
 tcgattaaagatattgaaacagtaaaacgaactgattttgaagaaatagatcagtttgat
 cctaaacgctttacgctaaataaagctaaaaaggatatcattactgatgttaataactcaa
 atccaaaatggtttcaagaaattgaaacaataaaagggtttaacttctaataaaaaact
 cagtttgataaacaattaactgcactacaaaaagaatttttagaaaaagtcgagcatgct
 50 catcaatttagtagaattaaatcaattacaacaagagtttaataatagatatgaacatt
 ttaaaacagcacatttactaggtgaaaaacatatagcagaacataaattaggatctgtt
 gtagtaaaacaaactcagcaataactaaataatcaatctgcttcttactttataaaacaa
 tgggcacttgatagaattaaacaaaattcaactagaacgatgaattcaattcgtggtgcg
 cataccgtaca

55

Sequence 2964

MGLTKSLVAKQPTVQKTSVYINEDQPEQSAYNDSITMGQTIINKTADPVLDKTLVDNAIS
 NISTKENALHGEQKLTTAKTEAINALNLTADLNTPQKEAIKTAINTAHTRTDVTAEQSKA
 NQINSAMHTLRQNISDNESVTNESNYINAEPEKQHFTEALNNAKEIVNEQQATLDANSI

NQKAQAILTTKNALDGEEQLRRAKENADQEINTLNQLTDAQRNSEKGLVNSSQTRTEVAS
 QLAKAKELNKVMEQLNNLINGKNQMINSSKFINEANQQQAYSNAIASAEVLKNKSNPE
 LDKVTIEQAINNINSAINNLNGEAKLTAKEDAVASINNLSGLTNEQKTENQAVVGSQT
 RDQVANVLRDSKALDQSMQTLRLDLVNNQNVIHSTSNYFNEDSTQKNYDNAIDNGSTYIT
 5 GQHNSSELNKSTIDQTIISQINTAKNDLHGAEKLRDKGTANQEIGQLGYLNDPQKSAEESL
 VNGSNTRESEVEHLNEAKSLNNAMKQLRDKVAEKTNVKQSSDYINDSTEHRQGYDQALQE
 AENIINEIGNPTLNKSEIEQKLQQLTDAQNALQGSLLLEEAKNNAITEINKLTALNDAQR
 QKAIENVQAQQTIPAVNQQLTLTREINTAMQALRDKVGQNNVHQSSNYFNEDEQPKHNY
 DNSVQAGQTIIDKLQDPIMNKNEIEQAINQINTTQTALSGENKLHTDQESTNRQIEGLSS
 10 LNTAQINAEKDLVNQAKTRTDVAQKLATAKEINSAMSNLRDGIQNKEDIKRSSAYINADP
 TKVTAYDQALQNAENIINATPNVELNKATIEQALSRVQQAQDLDGVQQLANAKQQAOTQ
 VNGLNSLNDGQKRELNLLINSANTRTKVQEELNKATESNHAMEALRNSVQNVQVKQSSN
 YVNEQPEQHNYDNAVNEAQATINNNAQPVLDKLAIERLTQTVNTTKDALHGTQKLIQDQ
 QAAETGIRGLTSLNEPQKNAEVAKVTAATTRDEVRNIRQEATTLDATMLGLRKS IKDKND
 15 TKNSSKYINEDHQQQAYDNAVNNAQHVIDETQATLSSDTINQLANAVTQAKSNLHGDTK
 LQHDKDSAKQTIQALQNLNSAQKHMEDSLIDNESTRTQVQHDLTEAQAALDGLMGALKESI
 KDNTNIVSNGNYINAEPSSKKQAYDAAVQNAQNIINGTNQPTINKGNVTTATQTVKNTKDA
 LDGDHRLEEAKNNANQTI RNLSNLNNAQKDAEKNLVNSASTLEQVQQLQTAQQLDNAMG
 ELRQSIANKDQVKADSKYLNEDPQIKQNYDDAVQRVETIINETQNPPELLKANIDQATQSV
 20 QNAEQALHGAEKLNQDKQTSSTELDGLTDLTDAQREKLREQINTSNSRDDIKQKIEQAKA
 LNDAMKKLKEQVAQKDGVBANSYDYNEDSAQKDAYNNALQAEDIINNSSNPNLNAQDIT
 NALNNIKQAQDNLHGAQKLQDQKNTTNQAIIGNLNLNQPQKDALIQAINGATSRDQVAEK
 LKEAEALDEAMQKLEDQVNVQDDQISNSSPFINESSDKQKTYNDKIQAAKEIINQTSNPTL
 DKQKIADTLQNIKDAVNNLHGDQKLAQSKQDANNQNLNHLDDLTEEQKNHFKPLINNADTR
 25 DEVNKKLEIAKQLNGDMSTLHKVINDKQDQIHLNSYINADNDKKQNYDNAIKEAEDLIHN
 HPDTLDHKAQDLDLNLINHAQTKQVQAEIIAQANKLNNEMGLTKTLVEEQSNVHQQSKYIN
 HVTTLLESLAQELQKAKELNDAMKAMRDSIMNQEIQIRKNSNYTNEDLAQQNAYNHAVDNIN
 NIIGEDNATMDPQIIKQATQDINTAINGLNGDQKLQDAKTDAKQQTINFTGLTEPQKQAL
 ENIINQOTSRANVAKQLSHAKFLNGKMEELKVAVAKASLVRQNSNYINEDVSEKEAYEQA
 30 IAKGQEIINSENNPTISSTDINRTIQEINDAEQNLHGENKLRQAQEI AKNEIQNL DGLNS
 AQITKLIQDIGRTTTKPAVTQKLEEAQAINQAMQQLKQSIADKDATLNSSSNYLNEDSEKK
 LAYDNAVSQAEQLINQLNDPTMDISNIQAITQKVIQAKDSLHGANKLAQNQADSNIINQ
 STNLNDKQKQALNDLINHAQTKQVQAEIIAQANKLNNEMGLTKTLVEEQSNVHQQSKYIN
 EDPQVQNIYNDISIQKGREILNGTTDDVLNNNKIADAIQNIHLTKNDLHGDQKLQKAQQDA
 35 TNELNYLTNLNNSQRQSEHDEINSAPSRTEVSNDLNHAKALNEAMRQLENEVALENSVKK
 LSDFINEDEAAQNEYSNALQKAKDIINGVPSSTLDKATIEDALLELQNARESLHGEQKLQ
 EAKNQAI AEIDNLQALNPGQVLAEKT LVNQASTKPEVQEALQKAKELNEAMKALKTEINK
 KEQIKADSRVYNADSLQANYNSALNYGSQIIATTQPPPELNKDVINRATQTIKTAENNLN
 GQSKLAEAKSDGNQSI EHLQGLTQSQKDKQHDLINQAQTKQVDDIVNNSKQLDNMSMNL
 40 QQIVNNNTVKNQSDFINESSQDAYNHAIQAADLITAHPTIMDKNQIQDAIENIKQA
 LNDLHGSNKLSQKKEASEQLQNLNLSLTNGQKDTILNHI FSAPTRSQVGEKIASAKQLNN
 TMKALRQSIADNNEILQSSKYFNEDESEQNAYNQAVNKAKNIINDQPTPVMANDEEQSVL
 NEVKQTKDNLHGDQKLANDKTDQATLNLALNYLNQAQRGNLETKVQNSNSRPEVQKVVQL
 ANQLNDAMKKLDDALTGNDAIKQTSNYINEDTSQQVNFDEYTDGRKNIVAEQTNPNMSPT
 45 NINTIADKITEAKNDLHGVQKLEQAQQQSINTINQMTGLNQAQKEQLNQEIQQTQTRSEV
 HQVINKAQALNDSMNTLRQSI TDEHEVKQTSNYINETVGNQTAYNNAVDRVKQIINQTSN
 PTMNPLEVERATSNVKT SKDALHGERELNDNKNKSTFAVNHLNDNLNQAQKEALTHEIEQA
 TIVSQVNNIYNKAKALNNDMKKLKDIVAQQDNVRQSNNYINEDSTPQNMYNNDTINHAQSI
 IDQVANPTMSHDEIENAINNIKHAINALDGEHKLQAKENANLLINSNLNAPQRDAIN
 50 RLVNEAQTRKVAEQQLQSAQALNDAMKHLRNSIQNQSSVRQESKYINASDAKKEQYNHAV
 REVENIINEQHPTLDKEI IKQLTDAVNQANNLNGVELLDADKQNAHQSIPTLMHLNQAQ
 QNALNEKINNNAVTRAKVAAIIGQAKILDHAMENLEESIKDKEQVKQSSNYINEDPDVQET
 YNNAVDHVTEILNQTVNPTLSIEDIEHAINEVNQAKKQLRGKQKLYQTIDLADKELSKLD
 DLTSSQSSSIQNIYTAKTREVAQAIEKAKSLNHAMKALNKIYKNADKVLDSRFINED
 55 QPEKEAYQQAINHVDISIHRQTNPEMDPTVINSITHELETAQNNLHGDQKLAHAKQDAAN
 VINGLIHLNVAQREVMINTNTNATREKVAKNLONAQALDKAMETLQQVVAHKNNILNDS
 KYLNEDSKYQQQYDRVIADAEQLLNQTTNPTLEPYKVDIVKDNVLANEKILFGAEKLSYD
 KSNANDEIKHMYNLNNAQKQSIKDMISHAALRTEVKQLLQQAKTLDAMKSLEDKTQVVI
 TDTTLPNYTEASEDKKEKVDQTVSHAQAIIDKINGSNVSLDQVRQALEQLTQASENLGDG

QRV E E A K V H A N Q T I D Q L T H L N S L Q Q T A K E S V K N A T K L E E I A T A S N N A L A L N K V M G K L E Q
 F I N H A D S I E N S D N Y R Q A D D D K I I A Y D D A L E H G Q D I Q K S N A T O N E A K Q A L Q Q L I N A E T S L N
 G F E R L N H A R P R A L E Y I K S L E K I N N A Q K S A L E D K V T Q S H O L L E L E H L V N E G T N L N D I M G E L
 A N A I V N N Y A P T K A S I N Y I N A D N L R K D N F T Q A I N N A R D A L N K T Q G Q N L D F N A I D T F K D D I F
 5 K T K D A L N G I E R L T A A K S K A E K L I D S L K F I N K A Q F T H A N D E I M N T N S I A Q L S R I V N Q A F D L
 N D A M K S L R D E L N N Q A F P V Q A S S N Y I N S D E D L K Q Q F D H A L S N A R K V L A K E N G K N L D E I Q I E
 G L K Q V I E D T K D A L N G I Q R L S K A K A K A I Q Y V Q S L S Y I N D A Q R H I A E S N I H N S D D L S S L A N T
 L S K A S D L D N A M K D L R D T L E S N S T S V P N S V N Y I N A D K N L Q I E F D E A L Q Q A S A T S S K T S E N P
 A T I E E V L G L S Q A I Y D T K N A L N G E Q R L A T E K S K D L K L I K G L K D L N K A Q L E D V T N K V N S A N T
 10 L T E L S Q L T Q S T L K L N D K M K L L R D K L K T L V N P V K A S L N Y R N A D Y N L K R Q F N K A L K E A K G V L
 N K N S G T N V N I N D I Q H L L T Q I D N A K D Q L N G E R R L K E H Q Q K S E V F I I K E L D I L N N A Q K A A I I
 N Q I R A S K D I K I I N Q I V D N A I E L N D A M Q G L K E H V A Q L T A T K D N I E Y L N A D E D L K I Q Y D Y A
 I N L A N N V L D K E N G T N K D A N I I I G M I Q N M D D A R A L L N G I E R L K D A Q T K A H N D I K D T L K R Q L
 D E I E H A N A T S N S K A Q A K Q M V N E E A R K A F S N I N H A T S N D L V N Q A K D E G Q S A I E H I H A D E L P
 15 K A K L D A N Q M I D Q K V E D I N H L I S Q N P N L S N E E K N K L I S Q I N K L V N G I K N E I Q Q A I N K Q Q I E
 N A T T K L D E V I E T T K K L I I A K A E A K Q V I K E L S Q K K R D A I N N N T D L T P S Q K A H A L A D I D K T E
 K D A L Q H I E N S N S I D D I N N N K E H A F N T L A H I I I W D T D Q Q P L V F E L P E L S L Q N A L V T S E V V V
 H R D E T I S L E S I I G A M T L T D E L K V N I V S L P N T D K V A D H L T A K V K V I L A D G S F V T V N V P V K V
 V E K E L Q I A K K D A I K T I D V L V K Q I K I D I S N N E L T S T Q R E D A K A E I E R L K K Q A I D K V T H S K
 20 S I K D I E T V K R T D F E E I D Q F D P K R F T L N K A K K D I I T D V N T Q I Q N G F K E I E T I K G L T S N E K T
 Q F D K Q L T A L Q K E F L E K V E H A H N L V E L N Q L Q Q E F N N R Y E H I L N Q A H L L G E K H I A E H K L G Y V
 V V N K T Q Q I L N N Q S A S Y F I K Q W A L D R I K Q I Q L E T M N S I R G A H T V X

Sequence 2965

25 Contig_0541_pos_1894_2919
 is similar to (with p-value 2.0e-28)
 >gp:gp|AB001577|AB001577_1 Pseudomonas sp. DNA for low spec
 ificity L-threonine aldolase, complete cds. NID: g2865133.
 gtgatttcatttgaaaatgattatttagaaggtgcacatgaaaaagtttaaatcgatta
 30 gtagagacaaatcgaatacaagctgctggatattggcttcgatgacttttcggcacaagct
 gcagataaaattagacaacgtattgactgtccagatgctaccattcgttttttagtaggt
 ggtacgcaaaccaatcaagtagttatttaactcaatgcttgatagttatgaaggtgttata
 tccgctgatacaggacatgtggcagtcctcatgaaggtggtgcgatagaattcagtgacat
 aaagtcttaaccataccctcccaagaaggttaagattactgctcaagacgcttgagaattat
 35 atagaaacttttgaaagtgatttttaaaaaagaacacatggtgtatccagggatggtttat
 atttcacatccaaccgaatatggaactttatacacgaaagaagaattacaatctttatct
 agagtttgccgtagacatcagattccactatttatggatggtgcacgcttaggctatggc
 cttatgagcaatcaactaatgtaactatcgaagatgttgcaaaatactgtgatgtgttt
 tacataggaggtactaagattggagcactttgtggtgaagcaattgtcttcactaaacaa
 40 aatgaacctataaaacttcactacaattataaaacatcatggtgctttatagcaaaaggc
 cgtctaactggtgttcaatttttagaattattcactgatgatttatattttgatataagt
 cgacatgctattaaaatggctgaaaaggtaaaaaaaggatttatagataaaggatatcaa
 gtctattttgatccaccaaccaatcaacaattttttattttaagcaacgataaaattgaa
 gaactaaaacaaaaggtaaaattcgagctttgggagaaatacgataatcaacatcgtgta
 45 gttcgcttcgcaacaagttgggcccacaactgaagaaaaatgttaatcaactacttgaacta
 atataa

Sequence 2966

50 VISFENDYLEGAHEKVLNRLVETNRIQAAGYGFDDFSAQAADKIRQRIDCPDATIRFLVG
 GTQTNQVVINSMLDSYEGVISADTGHVAVHEGGAIEFSGHKVLTIPSQEGKITAQDVENY
 IETFESDFKKEHVMVYPGMVYISHPTEYGTLYTKEELQSLSRVCRRHQIPLFMDGARLYG
 LMSNQTNVTIEDVAKYCDVFYIGGTKIGALCGEAIIVFTKQNEPKNFTTIKHHGALLAKG
 RLTVQFFLELFTDDLYFDIRHAIKMAEKVKKGFIKGYQVYFDSPTNQFFILSFQKIE
 55 ELKQKVFVAVWEKYDNQHRVVRFATSWATTEENVNQLLELI*

Sequence 2967

Contig_0541_pos_7422_6532
 is similar to (with p-value 3.0e-74)
 >sp:sp|P54204|FUR_STAEP FERRIC UPTAKE REGULATION PROTEIN. >

gp:gp|X97011|SEFURSOD_1 *S.epidermidis* genes fur and sod (partial). NID: g1263907.

5 atgttaataggatcacatgtttcaatgagtggtgcaaaaaaatgctgcaagggtcagcagaa
gaagcacataaatatggtgaatctacatttatgatttatacaggtgcgcctcaaaataca
agacgtaaaaatattgaagatttaaatatcgaaaaagccagcaggcaatgaaaacatat
10 ggcttatcaaatatcggtgtacatgcaccatataatcattaacattgcaaatacaaacaaa
cctgaagtatttaatttaggagtcgactttctacaaaaagaaatcgaaagaactcaagcg
ctcggagcgaagatatgttactgcatcctggagcgcgcatgtcggagcagggtgtagataaa
ggaattcaaaaaattattgaaggacttaatgaagtactcacacatgataatgatgtaaga
15 atagcacttgaaactatggcgggttaaaggaacagaagtagggagatcttttgaagaaatt
gctcaaataattgatggtgttacacataatgatcgcttatcagtatgttttgatacgtgc
cacactcatgatgcgggttataacgtcaaagaagatttcgatggtgtactagaaaaattc
gacagcataattggagtagatcgaattaaagtagtacatgttaatgacagtaaaaaaccta
agaggtgcacagaaggatcgtcacgaaaatatacggttttggtcatattggctttgatgca
20 ctttaattcgtagtagatcatgatacttttaaaaatattcccaaaatattagaaactcca
tatgttggtgaagataaaaaaaaataaaaaaccaccgtataaattagaaatagacatgtta
aaatcacaaaaatttgatccagaactcaaaaacaaaattttaactcaataa

Sequence 2968

20 MLIGSHVSMGSKMLQGSAAEEAHKYGESTFMIYTGAPQNTRRKNIEDLNIEKGQAMKTY
GLSNIVVHAPYIINIANTTKPEVFNLGVDLQKEIERTQALGAKDIVLHPGAHVAGVDK
GIQKIIIEGLNEVLTHDNDVRIALETMAGKGTEVGRSFEEIAQIIDGVTHNDRLSVCFDTC
HTHDAGYNVKEDFDGVLEKFDSSIIGVDRIKVVHVNDSKNLRGAQKDRHENIGFGHIGFDA
25 LNYVVHHDTEFNIPKILETPYVGEDKKNKKPPYKLEIDMLKSQKFDELPKLNKILTQ*

Sequence 2969

Contig_0541_pos_4647_4228

>sp:sp|P54476|END4_BACSU PROBABLE ENDONUCLEASE IV (EC 3.1.2
1.2) (ENDODEOXYRIBONUCLEASE IV). >gp:gp|D84432|BACJH642_144
30 *Bacillus subtilis* DNA, 283 Kb region containing skin element
. NID: g2627063. >gp:gp|299116|BSUB0013_223 *Bacillus subtilis*
s complete genome (section 13 of 21): from 2395261 to 261373
0. NID: g2634723.

atgaatacaaatgatgcaattaaagtttttaaaggaaaacggacttaaatatactgataaa
35 cgtaaagatatgctagatatctttgttaaagaggataaatattttaaagctaaacatatt
caacaacaaatggataaagactatcctggaatatcatttgatactgtatacagaaatctt
catttattttaaagatttaggcattatagagagtaccgaattagatggagaaatgaaattc
agaatcgcgatgcacaaatcatcaccatcatcattttatttgcgaaaattgaggagaaact
aaagtgtattgtttttgtccaatagaaaagattaaaagtcattaccgaatgtaaaatatt
40 catactcataaattagaagtgtatggtatttgtgaagaatgtcaacgtaaagcaaactaa

Sequence 2970

MNTNDAIKVLKENGLKYTDKRKMDLDFVKEDKYLNAKHIIQQQMDKDYPGISFDTVYRNL
45 HLFKDLGIIESTELDGEMKFRIACTNHHHHHFICENCGETKVIDFCPIEKIKSQLPNVNI
HTHKLEVYGICEECQRKAN*

Sequence 2971

Contig_0542_pos_5747_5274

is similar to (with p-value 2.0e-17)

50 >sp:sp|P41893|PPAL_SCHPO LOW MOLECULAR WEIGHT PHOSPHOTYROSINE
PROTEIN PHOSPHATASE (EC 3.1.3.48) (LOW MOLECULAR WEIGHT C
YTOSOLIC ACID PHOSPHATASE) (EC 3.1.3.2) (PTPASE) (SMALL TYRO
SINE PHOSPHATASE). >pir:pir|A55446|A55446 protein-tyrosine-p
55 hosphatase (EC 3.1.3.48), low molecular weight - fission yea
st (*Schizosaccharomyces pombe*) >gp:gp|L33929|YSPLMPTP_1 *Schi*
zosaccharomyces pombe low Mr protein tyrosine phosphatase mR
NA, complete cds. NID: g602991.

gtgataactaatgatacatgtagcatttgtatgtctcggtaatatatgtcgtttctccaatg
gctgaggctatcatgagacaaagactacaagaaagggtatttcagatatataaaagtcat

tctagaggaaacaggacgttggaatttaggcgaacctccacataacggaacacaaaaaatt
 ctacagaagtaccatattccttatgatggtatggtgagtgaaacttttcgaacctgatgat
 gattttgactatattattgctatggaccaaagtaacgtagacaatatcaaacaaatcaat
 ccaaatttacaaggacaattgttcaaattgctagaatttagtaacatggaagagagtgat
 5 gtaccagatccatactacacaaaataattttgaagggtgttttcgagatggtgcaatcatct
 tgtgataatttaatagactacatcgtaaaagatgcaaatttgaaagagaggttaa

Sequence 2972

VILMIHVAVFCLGNICRSPMAEAIMRQRLQERGISDIKVHSRGTGRWNLGEPHNGTQKI
 10 LQKYHIPYDGMVSELEFPDDDFDYIIAMDQSNVDNIKQINPNLQGQLFKLLEFSNMEESD
 VPDPPYTNFEGVFEMVQSSCDNLIDYIVKDLNLER*

Sequence 2973

Contig_0546_pos_1878_160

15 >sp:sp|P17444|BETA_ECOLI CHOLINE DEHYDROGENASE (EC 1.1.99.1
) (CHD). >gp:gp|X52905|ECBET_5 Escherichia coli betT, betI,
 betB and betA genes. NID: g48714. >gp:gp|M77738|ECOBETA_1 E.
 coli choline dehydrogenase (betA) gene, complete cds. NID: g
 145401. >gp:gp|AE000138|AE000138_2 Escherichia coli K-12 MG1
 20 655 section 28 of 400 of the complete genome. NID: g1786501.

atgagaagaaaacgcgattcatagcattatgtcatcattggtggcggtagtgaggttca
 gttcttgggtgcacgcctttcagaggataaagataaaaatgttttggatttagaagctgga
 cgtagtgactattttctgggatttatttattcaaatgccagcagcattgatgttcccatca
 25 ggtaatcggtttttatgactgggaatatcaaaactgacgaagaaccacatatgggacgtaga
 gtagatcatgacgagaggttaaggttaggtggctcaagttctattaacggtatgatttat
 caacgaggttaacccaatggactatgaaggatgggcagaacctgaaggaatggacacatgg
 gactttgcacattgtctaccatacttcaaaaagttagaaacaacatatggtgcagcgcca
 tacgataaaggttagaggccatgatgggtccaatcaaatataaacgtggaccagctactaat
 30 ccattattttaaatacattctttaatgcaggtgttgaagcggttatcataaaactgcagac
 gtttaatggatacagacaagaaggttttggaccatttgatagccaagtacatcatggacgt
 cgtatgtctgcttcaagagcgtatctacgcccagcattaagacgtagaaacttagatggt
 gaaacacgtgcattcgttacaaaattaatttttgatgaaaaataatagtaaaaaagtaaca
 35 ggcgtgactttcaagaaaaatggttaagaacatactgttcagcaaacgaagttatttta
 tctggcggtgctttcaatacaccacaactattacaattatcaggtattggtgactcagaa
 ttctttaaatacaaaaggtatagagccacgtatgcatttaccaggtgttgggtgagaacttc
 gaagatcacttagaagtatatattcaacataaatgtaaaacaaccgggtttcactacacact
 agccttgatgtcaaacgtatgccgttcacgtttacaatggatttttgcacgtaaaggt
 40 gcagtcgctgctcaaccactttgaagggtgggtgctttgtaagatcaaatgatgatgtgat
 tatccaaacccatcatgttccatttcttaccattgtctgtaagatatgatggtcaaaaagca
 ccagtagcacatggttaccaggtacatgttggaccaatgtactccaactcaagaggtagt
 ttgaaaatacaaatctaaggatccatttgaaaaaccaagtatcgtgtttaattacttatct
 acgaaagaagacgaaagagaatgggttgaaagcaattagagtagcaagaaatatcctaaaa
 caaaaagctatggaccatttaaatggtggcgaaatttcaccaggaccacaagttcaaacg
 45 gatgaagaaatttctagattgggtacgtaaagatggagaaactgcattacatccatcttgt
 agcgcgaaaaatgggacctgcatctgacccaatggcagtagtcgatccattaactatgaaa
 gtacatggtatggaaaatttacgtgtcgttgatgcttcagcaatgcctagaacaacaaat
 ggtaatatcatgcacctgtattgatgttagctgagaaagcagcggacattattcgtggt
 agaaaaccgcttgaaacctcaatatgttgactattataaacatggtattgatgatgaaaaa
 50 gcaggtgcaatggaagatgatccattctaccaatattaa

Sequence 2974

MRRKRDSYDYVIIGGGSAGSVLGRARLSEDKDKNVLVLEAGRSYFWDLFIQMPAALMFPS
 55 GNRFYDWEYQTDDEPHMGRVDHARGKVLGGSSSINGMIYQRGPNMDYEGWAEPEGMDTW
 DFAHCLPYFKKLETTYGAAPYDKVRGHDGPIKLKRGPATNPLFKSFFNAGVEAGYHKTD
 VN:YRQFEGFPGDSQVHHGRRMSASRAYLRPALRRRNLDVETRAFVTKLIFDENN?KKVT
 GVTFKKHGKEHTVHANEVILSGGAFNTPQLQLSGIGDSEFLKSKGIEPRMHLPG?GENF
 EDHLEVYIQHKCKQPVSLQPSLDVCRMFFIGLQWIFARKGAASNHFEGGGFVRSNDDVD
 YPNLMFHFELPIAVRYDGQKAPVAHGYQVHVGPMSNSRGSLSKIKSKDPFEKPSIVFNLYS

TKEDEREWEAIRVARNILKQKAMDPFNGGEISPGPQVQTDEEILDWVRKDGETALHPSC
SAKMGPASDEPMVVDPLTMKVHGMENLRVVDASAMPRTTNGNIHAPVLM LAEKAADIIRG
RKPLEPQYVDYYKHGIDDEKAGAMEDDPFYQY*

5 Sequence 2975

Contig_0547_pos_6940_6371

is similar to (with p-value 2.0e-34)

>pir:pir|S47148|S47148 hypothetical protein 1 - Staphylococcus carnosus >gp:gp|X79725|SCSECA_1 S.carnosus (TM300) secA gene. NID: g499333.

10 atgattagatttgaaattcatggagataacctcactatcacagatgcaattcgcaactat
attgaggagaaagtaggttaaattagaaagatactttaacaatgtgccaaatgctgtagca
catgttagagtgtaaaactatttctaattctacaactaaaattgaagttacaattccttta
aaagatgtcactcttagagctgaagaaagacatgatgatttatatgcaggcattgattta
15 attactaacaattggaaagacaagtacgtaaatacaaaactcgtgtaaactcgtaaacat
agagatcgtggagatcaagatatctttgttgctgaagtacaagagctctacaacaacaat
catgcagatgatatagaagtgaaaatgatattgaaattattcgttctaacaattcagc
ttaaaaccaattggattctgaagaagcagatttgcaaatgaatttattgggacatgatttc
ttatttttacogtagagaactgatggcacaagcattgtttatagacgaaaagatggc
20 aaatacggctctgattgaaacaactgaataa

Sequence 2976

MIRFEIHGDNLTITDAIRNYIEEKVGKLERVFNNVNAVAHVVRKTYNSSTTKIEVTIPL
KDVTLRAEERHDDL YAGIDLITNKLERQVRKYKTRVNRKHRDRGDDIFVAEVQESTTNN
25 HADDIESENDIEIIRSKQFSLKPMDSSEAVLQMNLLGHDFFIPTDRETGTSIVYRRKDG
KYGLIETTE*

Sequence 2977

Contig_0550_pos_2797_2369

is similar to (with p-value 5.0e-26)

>gp:gp|AF036166|AF036166_1 Xanthomonas campestris organic hydroperoxide resistance protein (ohr) gene, complete cds. NID: g3098341.

35 atggcaaatctctattttactcaactacaatgatttcaaatgggtggacgtgatggctcgctt
tttagtccagataatacatcttggttcaaaaccttgcaacaccttaaggaaatgggtggtaa
ggaggcaacgataactaatcctgaacaattatttgctgctggatagtgcatgctttaat
agcgcgctatcattaatcttatctcaaaataaaataagtgatgccaaaccagaagttaa
atcactattgaattacttaagatgatactgacaatgggttttaaaacttggcgagatatt
aaagtacacttgaaaatatgtcccaacaagatgctgagaaatttgaggagcaagcacat
40 caattctgtccatactcaaaagcgacacgtggttaacattgacgttcagttagatgttaca
gcgcaataa

Sequence 2978

MANSIYSTTMISNGGRDGRVFS PDNTFVQNLATPKEMGGQGGNDTNPEQLFAAGYSACFN
45 SALSLILSQNKISDANPEVEITIELLKDDTDNGFKLGADIKVTLENMSQQDAEKFVEQAH
QFCPYSKATRGNIDVQLDVTAQ*

Sequence 2979

Contig_0553_pos_7975_7067

is similar to (with p-value 1.0e-37)

>sp:sp|Q27546|IUNH_CRIFA INOSINE-URIDINE PREFERRING NUCLEOSIDE HYDROLASE (EC 3.2.2.1) (IU-NUCLEOSIDE HYDROLASE) (PURINE NUCLEOSIDASE).

55 atgtctattcccattataattgatactgatcctggatatagatgatgtacagcaattagt
atcgcactttcacacctcaatttgacgttaaaatgatatacaactgtgaatggtaatgta
ggtattgagaaaacgacagcaaatgcattaaagctaaaaaggttttttaatagtctgtt
cctgtacatagaggggcatcccaaccattgattaatgacatctttgaagctacatcaatt
catgggtgagctcgttatggatgggttacgagtttccacaaataaatcaagatgatttaaca
tcaattcatgcagttgaagcaatgagaaatctattagtaataactcaagaaccttaacc

ttgattgccataggtccactaacaatatcgctattcttttaactagttatcccgaagtt
 caaccattttattaaggaaattgttttaatgggtggttagtaccggtagaggtaatgtaacg
 ccttttagctgaatttaatatattgtgatccagaagctgctcaaattgtatttaactct
 ggattacctttgactatgattggctcttgatttggctcgtgaagcattgtttactcaccat
 5 tttgtaaaagacttcaaaagatacaaatgcaacttcaaacatgttatataatttatttcag
 cactataagagtgaagattttgaaataggatttaattatacgaatgtattcactatatta
 tatttggttgatccagaagcctttaatgtcaaggaggcgtataactcaaatagaattaaat
 ggcaactttacaaggggagccacagtggtagactttaatatggagcatcccaattgtaca
 10 gttgttttaagtctgttgaaagacagtatgaggatttattcttaaacgccctttcttat
 tgtaaataa

Sequence 2980

MSIPIIIDTPGIDDATAISIALSHPOFDVKMISTVNGNVGIEKTTANALKLRFFNSSV
 PVHRGASQPLINDIFEATSIHGESGMDGYEFPQINQDDLTSIHAVEAMRNLLVNTQEPLT
 15 LIAIGPLTNIAILLTSYPEVQPFKEIVLMGGSTGRGNVTPLAEFNIYCDPEAAQIVFNS
 GLPLTMIGLIDLAREALFTHHFKDFKDTNATSNMLYNLFQHYKSEDFEIGFKLYDVFTIL
 YLLDPEAFNVKEAYTQIELNGNFTRGATVVDNMEHPNCTVVLSPVERQYEDFLNALS
 YCK*

Sequence 2981

Contig_0554_pos_5578_0

is similar to (with p-value 9.0e-26)

>gp:gp|D87664|D87664_1 Thermus aquaticus DNA for DNA polyme
 25 rase family X, aminopeptidase T, QAH/OAS sulfhydrylase, comp
 lete cds. NID: g1526546.

atgacaaaaaaagatgtaattcaattattagaaaaaatagctatatatatggagctaaaa
 ggagaaaatacatttaaagtttcagcgtatagaaaagccgcacaaagtctagaggttgat
 gagcgtacattagaagagattgatgatgtaacagaacttaaggcattggaaaaggcgta
 ggagaagttattaatgaatttaaaacacaaggtcaatcatcgacccttcaagcacttcaa
 30 gatgaagtacctgaagggttagtgccacttttgaataacaaggattaggcagtaaaaaa
 atagcgaaactatcatgaacttcaaattacagataaagaaataacttcaaaaagcctgt
 gaagaaggtaagggtcagtcagtgctttaaagggtttgcaaaaaagacagagcaaaacatttta
 gaagcagtgaaagtcgatgggtgctaaaaaagatcggttatcctatagagctaatgagagga
 ctcaaccaagaaattgtaaaatttattgaacagttagaaggagttgaacaatttcaact
 35 gctggttagtttgcgaagatataaggaaatgagtaagatttagatttcataattagtaca
 tcagagcctaaaaaagttcaacaacaattacttcgtattccgaataaagtcaaagatggt
 gctggtggggataactaaaatttctctggaattagcttatgatgatgagacgattggcggt
 gatttttagattgatagaacctt

Sequence 2982

MTKKDVIQLEKIAIYMELKGENTFKVSAYRKAAQSLEVDERTLEEIDDVTELKGIGKV
 GEVINEFKTQGSSTLQALQDEVPEGLVPLLKIQGLGSKKIAKLYHELQITDKEILQKAC
 EEGKVSALKGFAKKTEQNILEAVKSMGAKKDRYPIELMRGLNQEIVKFIEQLEGVEQYST
 AGSFRRYKEMSKDLDFIISTSEPKKVQQQLLRIPNKVKDVAVGDTKISLELAYDDETIGV
 45 DFRLIEPX

Sequence 2983

Contig_0555_pos_3938_2001

>gp:gp|D85230|PEEGLTD_4 Plectonema boryanum URF141, ORF243,
 50 NADH-dependent glutamate synthase large subunit (gltB) and
 small subunit (gltD) and URF289 genes, partial and complete
 cds. NID: g1339947.

atgagttacggctctatctcagcagaggcacatgagacggttggtcaagctatgaatcaa
 attggaggtaaaagtaatagtggaaggtggagaagattcttcacggttacgaaattcaa
 55 aaggatggaagtaataagataagtgcgattaagcaagttgcatcaggtcggttttgggtg
 acgagtgattacttgcaacatgcaaaagaaattcaaattaaagtcgcacaaggcgctaaa
 ccaggggaaggtggacaactaccaggttcaaaagtatatccatggattgctgagactaga
 ggttcgacaccaggtataggattaatttcaccaccaccacaccatgatatttattcaatt
 gaggaattagcacagctcattcatgatttaaaaaatgcaaatagaagagctgataatgca

gttaagcttgatcaaaaactggcgttggaaactatagcttcaggggtagctaaagctttc
 gccgataaaattgttataagtggtatgatggaggtacaggtgcatcgctaaaacaagt
 attcaacatgcagggttgccatgggagataggccttgccgaaacacatcaaacacttaaa
 taaatgatttgcgtagtcgcgtaaaattagaaacggatggtaagttactgacgggtaaa
 5 gatgtagcttatgcttgtgcgcttggtgcagaagaatttggttcgcaacagcaccactt
 gttgttttgggtgtattatgatgaggggttgtcataacgatacgtgtccagtaggggtt
 gcaacacaaaacaaagatttaagagcttgttttagaggttaaggcacagcatgtagttaac
 tttatgtattttatagctgaagaattacgtgaaattttggcttcaacttggtttagaaaca
 gtagaagagtttagtaggaagaacagatcttcttcaacggttcgacgcaattgaaaccaa
 10 agtaagcagcttcgcttcaaatagaacggtttaatagaacaatttgacgggggttaatac
 gaagagatatcaaaaaccatcatcttgatgaaggattcgatttgaattatctgtacca
 gacgcacgctatagtattgaaaacgggcactctttaccggaaattatgttggttaata
 gaacagcgagatgtaggtgtaattacaggtagtcgataagtaacaatatggagaaga
 ggattacctgaagatacagatacttgccttactgaaggatcatgcaggtcaaagcttagct
 15 gcatatgcaccacgcggattaacaatccatcataccggtgatgctaatactacgttaggt
 aaaggattgtccggtggaactgtcatcgtaaatgctccaaatagtcaacgtgaaaatgaa
 attatagcaggaataataaacttttaccgggcttctagaggtaaagcgtttatcaatggt
 aaagctggtgagcgtttctgtatcagaatagtgtgagatggttagtagaaggtatt
 ggtgatcatggacttgaatatatgacagggggacatgtcattatcttaggagatggttga
 20 aagaactttggccaaggcatgagcggggcgtaagtatatatttctctctgacgtgga
 aaatttaaaaagggttaatgagcgttgaactttagaattcagtagcatacgttttgatgag
 gaaaaatctcttatcaaagacatgcttgaagcacattttaagcatacacgtagtaacaaa
 gcacgccaattacttgaccaatttgacaatattgaaaagtttagcaattaaagttattccg
 aaagattacaaattaatgatgcaaaaaattgatttgaaaaaacgtcaaaggaacgtgaa
 25 gatgaagcaacactggcagcgttttatgatgacagagaaacaattgaacaagagctacag
 ccagcagtcatttattaa

Sequence 2984

MSYGSISAEAHETLAQAMNQIGGKSNSGEGGEDSSRYEIQKDGSNKISAIKQVASGRFGV
 30 TSDYLQHAKEIQIKVAQGAKEGGQLPGSKVYPWIAETRGSTPGIGLISPPPHDIYSI
 EDIAQLJHDLKNANRRADIIVKLVSCTGVGTIASGVAKAFADKIVISGYDGGTGASPKTS
 IQHAGLFWIEGLAETHQTLKLNDRSRVKLETGKLLTGKDVAYACALGAEEFGFATAPL
 VVLGCIMMRVCHNDTCFVGATQNKDLRALFRGKAQHVVNFMYFIAEELREILASLGLET
 VEELVGRDQLQRSTQLKPNKAASLQIERLIEQFDGVNTKEISQNHHLDEGFDLNYLYP
 35 DARYSIENGHSFTGNVYVNNQRDVGVITGSAIAKQYGEGLPEDTILAYTEGHAGQSLA
 AYAPRGLTIHHTGDANDYVVGKLSGGTVIVNAPNSQRENEIIAGNINFYGASRGKAFING
 KAGERFCIRNSGADVVEGIGDHGLEMTGGHVIIIGDVGKNFGQMSGGVSYIFSSDVE
 KFKKVNALETLEFSSIRFDEEKSLEKIDMLEAHFKHTRSNNKARQLLDQFDNIEKLAIKVIP
 40 KDYKLMMQKIDLKKRQMEREDEATLAAFYDDRETIEQELQPAVIY*

Sequence 2985

Contig_0555_pos_1962_520

>sp:sp|P39812|GLTB_BACSU GLUTAMATE SYNTHASE [NADPH] LARGE C
 HAIN (EC 1.4.1.13) (NADPH-GOGAT). >gp:gp|Z99113|BSUB0010_138
 45 Bacillus subtilis complete genome (section 10 of 21): from
 1781201 to 2014980. NID: g2634090. >gp:gp|Z99114|BSUB0011_9
 Bacillus subtilis complete genome (section 11 of 21): from 2
 000171 to 2207900. NID: g2634230.
 atgaaatatgataaacagtcgctatcagaattgtctttgtagaccgtctttcgaatcat
 50 gaagcgtttcaacaacgcttcaactaaagaagatgcttcgattcagggtagcgcgctatg
 gacgcggaacacctttttgtcaaaactgggaatcttatggaagagaaacaatag;atgc
 cctattggtaattatatacctgagtggaacgacttagtctatcatcaagattttaagct
 gcttacgaaagattgagagagacgaataatttctgaatttacaggaagagtttgcct
 gcaccatgtagcaatcatgtgttatgaaaattaatagagaatccgtggcgattaaaggt
 55 attgaacgtgacaattattgatgaagcatatgagaatgagtggttcatcccgcataatcct
 gaagatcataaagaccaacgagttgctatcgtaggtagtggtccagcgggacttacagca
 gctgaagaattaaactttaaaggctataaagttactgtttatgaaaaggcgcatgaacca
 ggcggcttgctaattgtatggtataccaaatatgaaactagataaagacgtaatacgtcga
 cgtgtatcacttatgaaagatgctgggggtttattttaaacaggcgttgaaattggcgtc

gatgtgagccgtgaaacacttgaagaaaattatgatgctattatatttatgcacaggtgct
 caaaatgcgagagatttaccattggaaggacgaatgggctctggtattcattttgcaatg
 gactatcttactgaacaaacacagtatctaattgggtgagattgaaagtttgagca"tact
 gctaaagataagaatgtaattattataggtgctgggtgatactgggtgcagactgtgtagcg
 5 acagcattacgtgaaaactgtaaatctattgttcaatttaataaataacgaaacagcct
 gaagagattacttttgaaagtaatacttcctggccattagcaatgcctgttttcaaaatg
 gattatgcgcataaggaatatgaagctaaatttgggtcaagaaccaagagcctatgggtga
 caaacaatgcgctatgatgttgacgagtttaggaaatgttaaaggcttatatacacaata
 ttaaaagaacgcctgatggcatgggtgatggaagatggaccagaacgattttggccggct
 10 gatttagtcttattatctataggggttgggtggtactgaaaccactgttccgcatgcgttt
 gatatacacaccgagcgtaataaaattgtagctaataatacaaatatcaaaactaatcac
 gctaaaatatttgcgtgcaggagatgcaagacgaggtcagagtttgggtgttgggcaata
 aaagaaggtcgtgaagtagcacattctgttgatcaataacttaagtaagaagttctagtg
 taa

15

Sequence 2986

MKYDKQSLSELSLVDRLSNHEAFQQRFTKEDASIQGARCMDCGTPFCQTGQSYGREITGC
 PIGNYIPEWNLVYHQDFKAAAYERLRETNNFPEFTGRVCPAPCEQSCVMKINRESVAIKG
 IERTIIIDEAYENWVHPAYPEDHKDQ RVAIVGSGPAGLTAAEELNFKGYKVTVYEKAHEP
 20 GGLLMYGIENMKLDKDVIRRRVSLMKDAGVLFKTGVEIGVDVSRETLEENYDAIILCTGA
 QNARDLPLEGRMGSGIHFAMDYLTEQTQYLNGEIESLSITAKDKNVIIIGAGDTGADCVA
 TALRENCKSIVQFNKYTKQPEEITFESNTSWPLAMPVFKMDYAHKEYEAKFGQEPRA YGV
 QTMRYDVBELGNVKGlyTQILKETPDGMVMEDGPERFWPADLVLSIGFVGTETTVPHAF
 DIHTERNKIVANDTNYQTNHAKIFAAGDARRGQSLVVWAIKEGREVAHSVDQYLSKEVLV
 25 *

Sequence 2987

Contig_0557_pos_8699_9661

is similar to (with p-value 1.0e-71)

30 >sp:sp|P76113|YNCB_ECOLI PUTATIVE NADP-DEPENDENT OXIDOREDUC
 TASE IN TEHB-RHSE INTERGENIC REGION (EC 1.-.-.-). >gp:gp|D90
 784|D90784_8 E.coli genomic DNA, Kohara clone #273(32.5-32.8
 min.). NID: g1742353. >gp:gp|D90785|D90785_2 E.coli genomic
 DNA, Kohara clone #274(32.7-33.0 min.). NID: g1742363.
 35 atgccacaagacgatacattttaaataatgaagatatagatgttattgaaccttcagaaaat
 gaattgcaattgaaaacattatataatcgggtgatccatatatgagaggacgtatgaca
 aatgctgattcttatgtagatccggttcaaacaaaggggaaccgttcaatggacatacggta
 tctaaagttttgaaatccaaggatagtaattttgatgaaggtgatatagtagtgggtatg
 ctaccttggaagaaaaataaacagtaaatagttagtattgttaacaaagtacacttct
 40 gacattaccattacattcttcttagtgtgttggggatgcctgggtcagactgcttatcat
 ggattacttgatattggacaacctaaagaaggtgaaacagtagttatttcagcagcttca
 ggtgacagtgggttcagttgtgggcaaatgcaagcttaaaaggtgcagagtcg"tgggt
 atagctggtggagataaaaaagtgaaactatctaaaaaatgaacttcgttttgatgtgggt
 atcgattacaaaaaagataatttccctgaagcgttaaaagaagcgggtgcctaacggtata
 45 gatgtctacttcgaaaatgtaggtggatatattggcgatgaagcttcaaacatctcaat
 acacatgcaagaattcctgtttgtggtgcatctcctataatcatccagaaaaagat
 attggaccacgcattcagcaaacattgattaaaaatcaagcaatgatgagaggtttcata
 gtagcagaattcgctgatggttttaagaagcgagcaaacattagctcaatgggttcaa
 gagaataaaattaaaacacaagtttcagtagaagatgggtttgataaagtgccgcaagcc
 50 ttagaaaatctgctaactggtgataattttggttaacaagttattaaagtggcaagtgaat
 taa

Sequence 2988

MPQDDTFKYEDIDVIEPSENELQLKTLYISVDPYMRGRMTNADSYVDPFKQGEFPNGHTV
 55 SKVLKSKDSNFDEGDIVGMLPWRKINTVNSEYVNVKVPSTSDVPLHLYLSVLGMPGQTAYH
 GLLDIGQPKEGETTVVISASGAVGSVVGQIAKLKGRVVGIIAGGDKKVNLYLKNELRFDAG
 IDYKKDNFPEALKEAVPNGIDVYFENVGGYIGDEVFKHLNTHARIPVCGAISSYNHPEKD
 IGPRIQQTLIKNQAMMRGFIVAEFADGFKASKQLAQWVQENKIKTQVSVEDGFDKVPQA
 FRNLLTGDNFGKQVIKVASE*

Sequence 2989

Contig_0557_pos_9913_0

is similar to (with p-value 3.0e-17)

- 5 >sp:sp|Q08257|QOR_HUMAN QUINONE OXIDOREDUCTASE (EC 1.6.5.5)
(NADPH:QUINONE REDUCTASE) (ZETA- CRYSTALLIN). >pir:pir|PN04
48|PN0448 zeta-crystallin / quinone reductase (NADPH) (EC 1.
6.-.-) - human >gp:gp|L13278|HUMQUINZ_1 Homo sapiens zeta-cr
ystallin/quinone reductase mRNA, complete cds. NID: g292414.

10 atgccagttgacaaagcgccacgtgtacttggctttgatgctgttggtgtgattgaaaag
ataggagatcaagtgtcaatgtttcaagaaggggacgtcggtttttattcaggttctcct
aaccaaaatggttcgaatgaagaataccaattaatagaggaatatttagtagctaaagca
cctacaaatttgaaaagtgaacaagcagctagcctacctttaactgggetaacagcttat
15 gaaacgcttttcgatgtttttggaatttcaaaagaaccatctgaaaataaaggtaaataca
ttgttaataattaatggagcaggtggtgtaggtagtattgcaacacagatagcgaaattt
tatgggttgaaggttattacaactgcttcgagagaggatactataaagtggtctgttaat
atgggtgctgatgtgtgtactgaatcataaagaagatttaagtc

20 Sequence 2990

MPVDKAPRVLGFDAVGVIKIGDQVSMFQEGDVVFYSGSPNQNGSNEEYQLIEEYLVAKA
PTNLKSEQAASLPLTGLTAYETLFDVFGISKEPSENKGSLLIINGAGGVGSIATQIAKF
YGLKVITTASREDTIKWSVNMGADVVLNHHKDLX

25 Sequence 2991

Contig_0558_pos_11954_0

is similar to (with p-value 1.0e-22)

- >pir:pir|A43577|A43577 regulatory protein pfor - Clostridium
m perfringens
- 30 gtgttatgggagtggttgatattgatttgccttataggtacattattttaatttttagtt
ttagttgtttttactctatttacatataaggcacctagtgggatgaaagcgatgggtgct
ttagcgaatgcgcaattgcttcgttttttagttgaagctttcaataaataatgtgggtgga
caagtatttggtatttaattcttagaagaattagagatgctgctggaggtttaggtggt
gtggctgcggctggattaactgcattagcgattggagtatctccagtttatgcactagtt
35 attggtgcagcttgaggaggtatggacttgttgccaggatttttcgctgggtatattgta
ggctacatgatgaagtataccgagaatatgtgccagatggtattgatttaattg

Sequence 2992

- 40 VLWGVFDMDLLIGTLFLILVLVFTLFTYKAPSGMKAMGALANAIIASFLVEAFNKYVGG
QVFGIKFLEELGDAAGGLGGVAAAGLTALAIGVSPVYALVIGAACGGMDLLPGFFAGYIV
GYMMKYTEKYVPDGIDLIX

Sequence 2993

Contig_0558_pos_7113_6400

- 45 is similar to (with p-value 2.0e-21)

- >pir:pir|A55345|A55345 diamine N-acetyltransferase (EC 2.3.
1.57) - Escherichia coli >gp:gp|D25276|ECOSN1A_1 Escherichia
coli gene for spermidine acetyltransferase, complete cds. N
ID: g517104. >gp:gp|AE000254|AE000254_5 Escherichia coli K-1
50 2 MG1655 section 144 of 400 of the complete genome. NID: g17
87862.

atgagtcacttcgaacattttactttagaacattaccactccctttttcataatgatcgg
ttgatgtcgggtgatttttaatacagttgcagtgccacttttatctgcacgattgcaaca
gttaraggtacattttggtgcgatcgctttatattacttacgcaacaacgggttaagggtt
55 acgctattaacgatgaataatgtattaatggtatcttcagacgtagtcattgggtgcttct
tttttaattatgttcactgcaataggacatttcacagggttaggttaggattttcaaca
gtgctagcatctcatatagcggttctgtattccaattggtgttatcatcgcttaccctaa
ttatatgaaatgaatgataacatgttaaatgcagcaagagatttaggtgctaataatgaatca
caattattaactagcattattatacctaataattatgccctcgattataggaggattcttt

atggcattaacatattctctagatgattttacggtaagcttctttgttactggaaacgga
 tttagtgtattgtctgtagaagtttatgctatggctcgaaaagggtataagtatggagatt
 aatgccatttctacaataatattttattgctattatgtttggagtatttggatattacttt
 attcaacatatcgtaatacgtcagaaaaagatgaagcgagggtgtaataatgaatga

5

Sequence 2994

MSHFEHFTLEHYHSLFHNDRLMSVIFNTVAVALLSASIAITVIGTFGAIALYYLRNKRFKV
 TLLTMNNVLMVSSDVVIGASFLIMFTAIGHFTGLGLGFSTVLASHIAFCIPIVVIIVLPQ
 LYEMNDNMLNAARDLGANESQLLTSIIIPNIMPSIIGGFFMALTYSLDDFTVSFFVTGNG
 FSVLSVEVYAMARKGISMEINAISTIIIFIAIMFGVFGYYFIQHIVNRQKKMKRGVNE*

10

Sequence 2995

Contig_0558_pos_1973_1638

is similar to (with p-value 1.0e-26)

15

>sp:sp|P23859|POTC_ECOLI SPERMIDINE/PUTRESCINE TRANSPORT SY
 STEM PERMEASE PROTEIN POTC. >pir:pir|C40840|C40840 spermidin
 e/putrescine transmembrane protein C - Escherichia coli >gp:
 gp|D90747|D90747_3 Escherichia coli genomic DNA. (25.1 - 25.5
 min). NID: g1651548. >gp:gp|AE000212|AE000212_10 Escherichi
 a coli K-12 MG1655 section 102 of 400 of the complete genome
 . NID: g1787358. >gp:gp|M64519|ECOPOTABCD_3 E.coli transport
 protein (potA, potB, potC and potD) genes, complete cds. NI
 D: g147325.

20

gtggtagagtattagaaattaactttatacatagaacttgtgaagtgttaattattatc
 gatccgcagtatgcaaataatgggtacgcgaaaaagcctttaaaatggctattgactat
 gcttttttagtattaaatatgaataaggtatacttatatgtggatattaagaatgagaaa
 gcagtcacatatctatcaaaagtaataatttcgaaatagaaggaaacgttaaaaggaacacttc
 tatacaaggggagaatatagagattgctatgtaatgggcttgtaaaaaggaattggggt
 aataagaatgatgatgatttgcctcatataagatga

30

Sequence 2996

VVELLEINFHRTCEVLIIIDPQYANNGYAKKAFKMAIDYAFVLNMMNKVLYVDIKNEK
 AVHIYQSNFEIEGTLKEHFYTRGEYRDCYVMGLLKRNWVNKNDDDLSHIR*

35

Sequence 2997

Contig_0561_pos_2326_3213

is similar to (with p-value 4.0e-30)

>sp:sp|P28246|BCR_ECOLI BICYCLOMYCIN RESISTANCE PROTEIN (SU
 LFONAMIDE RESISTANCE PROTEIN).

40

atgatgatgactacaagttccaaacactttatctaaaatattaatcgttatacttggcggt
 atgactgcatttggctccttggactattgatatgtacggaccatctttacctaagttcag
 catgcgttgggttcacatcaatttcagaaatacaacttacattatcctttgctatgatagg
 ttagctattgggtcaatttgcattcgccactatcagatgtattaggtcgtaaaaaaatg
 gcactcattttattgattggatatttaatagcctcattattatcagtttttacagttcat
 ttaaaatatttttaattatccggtttaattcaaggttagcaggagggtgggtgcaatcgtc
 atagccaaagcttctattggagataactatgacggagacgaattagcaaaatttttaact
 tctcttatgggtcataaacggtataatcaccatcattgctccactgttaggtggcctcgct
 ttatctattgcaagttggagaatgatttttatatttttaacaatcattaccttaatagtc
 atcttaggcattttattaaagatgccagttgggccccatcaagaacaatctcagttaaat
 tttaaagcaatatttaagattttgggtctgttattaacaaaaccaccttcggtattcca
 atgttattgcaaggattaaacttatgtcatgttattcagttattcgtcagccgcacctttt
 atttcacaaaagatgtatcatatgacaccacttcaatacagtgcaatggttgctattaat
 ggagtgggttgatagtcgctcagtcagataaccgctattatagtagaaaaggtaagccga
 tatgcgatgctcatatatttaacaatcattcaaatgttaggtgtttaa

55

Sequence 2998

MMMTTSSKHLISKILIVILGVMTAFGPLTIDMYGPSLPKVQHAFGSSISEIQLTSLFAMIG
 LAIGQFVFGPLSDVLGRKKMALILLIGYLIASLLSVFTVHLTIFLIIRLIQGLAGGGAIV
 IAKASIGDNYDDELAKFLTSLMvingiitiiaplglgllalsiaswrmiifltiitliv

ILGILLKMPVGP HQEQSQLNFKAIFKDFGLLLTKPTFVIPMLLQGLTYVMLFSYSSAAPF
ISQKMYHMTPLQYSAMFAINGVGLIVVSQITAIIVEKVSRYAMLIYLTIIQMLGV*

Sequence 2999

5 Contig_0561_pos_4708_4139
is similar to (with p-value 3.0e-49)
>sp:sp|P28368|YVYD_BACSU HYPOTHETICAL 22.0 KD PROTEIN IN FL
IT-SECA INTERGENIC REGION. >gp:gp|Z31376|BSFLIDST_5 B.subtil
is (HB2058) genes for Flid, FliS, FliT proteins. NID: g49937
10 9. >gp:gp|Z99122|BSUB0019_28 Bacillus subtilis complete geno
me (section 19 of 21): from 3597091 to 3809700. NID: g263602
9. >gp:gp|U56901|BSU56901_22 Bacillus subtilis putative tran
scriptional regulator (yvhJ), Ycr59c/YigZ homolog (yvhK), hi
15 stidine kinase (degS), transcriptional regulator of degradati
on enzyme (degU), (degV), (comFA), (comFB), (comFC), flagell
ar protein (yviB), negative regulator of flagellin (flgM), f
lagellar protein (yviC), flagellar-hook associated protein 1
(flgK), flagellar-hook associated protein 3 (flgL), (yviE),
20 transmembrane protein (yviF), (csrA), flagellin (hag), flag
ellar protein (yviH), flagellar hook-associated protein 2 (f
liD), flagellar protein (fliS), flagellar protein (fliT), si
gma-54 modulator homolog (yviI), and (secA) genes, complete
cds. NID: gl762326.
atgattagatttgaaattcatggagataacctcactatcacagatgcaattcgcaactat
25 attgaggagaaagtaggttaaattagaaagatactttaacaatgtgccaatgctgtagca
catgttagagttaaaacttattctaattctacaactaaaattgaagttacaattccttta
aaagatgtcactcttagagctgaagaaagacatgatgatttatatgcaggcattgattta
attactaacaattggaaagacaagtagtaatacaaaactcgtgtaaactcgtaaactat
agagatcgtggagatcaagatatctttgttgctgaagtacaagagcttacaacaacaat
30 catgcagatgatatagaaagtgaatgatattgaaattattcgttctaaacaattcagc
ttaaaccatggattctgaagaagcagattgcaaattgaatttattgggacatgatttc
tttattttaccgatagagaaactgatggcacaagcattgtttatagacgaaaagatggc
aaatacggctctgattgaaacaactgaataa

35 Sequence 3000
MIRFEIHGDNLTITDAIRNYIEEKVGKLERVFNNVNPNAVAHVVRVKTYSNSTTKIEVTIPL
KDVTFRAERHDDLTYAGIDLITNKLERQVRKYKTRVNRKHRDRGDDQDIFVAEVQESTNN
HADDIESENDIEIRSKQFSLKPMDSSEAVLQMNLLGHDFIFTDRETDGTSIVYRRKDG
40 KYGLIETTE*

Sequence 3001
Contig_0562_pos_12885_13190
is similar to (with p-value 1.0e-20)
>sp:sp|P46378|FAS6_RHOFA HYPOTHETICAL 21.1 KD PROTEIN IN FA
45 SCIATION LOCUS (ORF6). >pir:pir|F55578|F55578 hypothetical p
rotein 2 (ipt 3' region) - Rhodococcus fascians plasmid pFiD
188 >gp:gp|Z29635|RFCCIPTFD_6 R.fascians (D188) genes for P4
50 cytochrome, isopentenyltransferase and ferredoxine. NID:
g455000.
gtgttgaaaaagatacatgggttattgttttgtccaactcttgaccatgatacaacttcc
gttatttcttctaaagttccaggtccaccaggaagtgtcatgcatacatcgcttttctt
aagataacttcttttctttcagacatactttccacgattattagttcgttctaatttatca
tgtgctaactctctttgtttttaaaccgttaggcattacaccaattgttttccggttgta
55 tgtattacagtatattgcaattgttcccattaatcctgcatttccgccacaaaaactaaa
gtgtga

Sequence 3002
VLKKIHGLLFCPTLDHDTTSVISSKVPGPSAMHTSPFFKITSFLSDILSTIISSSNLS
CANSLCFKNVGITPIVFLLCITVFAIVPINPAFPPPKTKV*

Sequence 3003

Contig_0562_pos_13285_12725

is similar to (with p-value 8.0e-24)

5 >sp:sp|P46378|FAS6_RHOFA HYPOTHETICAL 21.1 KD PROTEIN IN FA
SCIATION LOCUS (ORF6). >pir:pir|F55578|F55578 hypothetical p
rotein 2 (ipt 3' region) - Rhodococcus fascians plasmid pFiD
188 >gp:gp|Z29635|RFCCIPTFD_6 R.fascians (D188) genes for P4
50 cytochrome, isopentenyltransferase and ferredoxine. NID:
10 g455000.
atgaatattatagctctattgtggagcaagcaagggaaacaaaaaagaatacgaaaatagt
gcaattcgaattaggtgaatggatagctaaaaataatcacactttagtttttggtggcgga
aatgcaggattaatgggaacaattgcaaatactgtaatacataacaacggaaaaacaatt
gggtgaatgcctacgttttttaaaacaaagagagtttagcacatgataaattagacgaacta
15 ataatcgtggaaagtatgtctgaaagaaaagaagttatcttaagaaaggcgatgtatgc
atagcacttcctgggtggacctggaacttttagaagaaataacggaagttgtatcatggtca
agagttggacaaaacaataacccatgtatcttttcaacacaaataattattattccctt
atcgaacaattctacgatcaaattggtttcaaacgagtttttaactcaagaagatagagat
aaaatattattctcaaactcattccaagaaattgaagaatttatagaaaactataaaaca
20 caaaaataagaacttattaa

Sequence 3004

MNIIVYCGASKGNKKEYENSAIQLGEWIAKNNHTLVFGGGNAGLMGTIANTVIHNNKTI
GVMPTFLKQRELAHDKLDELIIVESMSERKEVILKKGDVCIALPGGPGTLEEITEVVSWS
25 RVGQNNNPCIFFNTNNYSLIEQFYDQMVSNFLTQEDRDKILFSNSFQEIEEFIENYKT
PKIRTY*

Sequence 3005

Contig_0563_pos_5497_5093

is similar to (with p-value 4.0e-71)

30 >gp:gp|AF046871|AF046871_3 Anabaena PCC7120 heterocyst-inhi
biting signaling peptide (patS) and holiday junction resolu
se (ruvC) genes, complete cds; and unknown genes. NID: g2896
023.
35 atgttagctaataatggtttaatcgcgattaatctcgcttatcagaatttagaaagagca
tttgttcaagatgtttctgatattgaatccaaacttacgttagcagcgacacctaagctc
gcatcaaaatcagctattagagaaagtatacgcttagcaattgttcctacaattgattct
gtaaaaacatatggtctagtttcaattccaggtatgatgacaggattgattatcggaggc
gttgaccacttcaagcaattaaatttcaattgcttgctggtttattcatacaacagcg
40 acgattatgtctgcactcattgcaacgtatatgagttacggtcagttccttaatgctcgt
catcaactcattgctagaacgcaacgcacaaagacaaagtagttaa

Sequence 3006

MLANGLIAINLAYQNLERAFVQDVSDIESKLTLAATPKLASKSAIRESIRLAIVPTIDS
45 VKTYGLVSIPGMMTGLIIGGVDPLQAIKFQLLVFIHTTATIMSALIATYMSYGQFFNAR
HQLIARTQRTQSS*

Sequence 3007

Contig_0563_pos_1994_1008

is similar to (with p-value 2.0e-32)

50 >sp:sp|P77307|YBBM_ECOLI HYPOTHETICAL 28.2 KD PROTEIN IN US
HA-TESA INTERGENIC REGION.
atgaatgtagtacttattggtgggtggcactggactttctgtccttgctagaggccttaga
gaatttccaatagacattactgccattgttactgtacggacaatgggtgggagcacgggg
55 aaaattagagatgtcatggatattccagcgccctggtgatattcgtaattgtcattgctgct
ttaagtgactcagaatcgatattaactcaattgttccagtaccgttttggtgaaaatcaa
gtagatgggcattcattaggtaatatttagttattgctggaatgactaacattactaatgat
tttgacacgctattaagagagttaagcaaaagttttaaatattaaggccaagtcacccct
tcaacaaacgcaagtgtgcaactcaacgcggtgatggaagacggtgaaattgtacatgga

gaaactaatatacctaaaacacataaaaaaatagatcgtgtgtttttagaaccaagtgtat
 gttgaaccaatgaatgaagcgatagaagcttttagaacaagcagatttaattgtccttagga
 ccagggttcattatatacaagtgttatatcaaatttatgtgtcaaaggatattcagaagca
 ttattacgtacatctgctccaaaactttatgtatctaattgttatgacacaaccaggcgag
 5 actgataattatgatgtcaaagagcatattgatgcacttactcgacaagttggtgaacca
 tttattgattttgtcatatgtagctcagaatcctatagtaaagatgttttacaacgatat
 gaagaaaagaattcgaaaccagtagcagtagacataaagaacaattaaaagatagtgaatt
 agagttttaacggcatctaatttagttgaaatatctaataaacactatgtcagacataac
 acaaaagtattatcaaaaatgatttatgaacttgccttagaattaacaagtacaattcgc
 10 tttactcctagtataaaaagaataa

Sequence 3008

MNVVLIGGGTGLSVLARGLREFPIDITAIVTVADNGGSTGKIRDVMDIPAPGDIRNVIAA
 LSDSESLTQLFQYRFGENQVDGHSGLNLVIAGMTNITNDFGHAIKELSKVLNIKQGVIP
 15 STNASVQLNAVMEDEIVHGETNIPKTHKKIDRVFLEPSDVEPMNEAIEALEQADLIVLG
 PGSLYTSVISNLCVKGISEALLRTSAPKLYVSNVMTQPGETDNYDVKEHIDALTRQVGEP
 FIDFVICSSSEYSKDVLRQYEEKNSKPVAVHKEQLKDSGIRVLTASNLEISNEHYVRHN
 TKVLSKMIYELALELTSTIRFTPSDKKK*

Sequence 3009

Contig_0564_pos_5882_6601
 is similar to (with p-value 7.0e-36)
 >sp:sp|P49309|MOCR_RHIME PROBABLE RHIZOPINE CATABOLISM REGU
 LATORY PROTEIN MOCR. >pir:pir|S51574|S51574 mocR protein - R
 25 hizobium meliloti >gp:gp|X78503|RMMOCCABR_6 R.meliloti mocC,
 ORF334, ORF293, mocA, mocB and mocR genes. NID: g468758.
 atgattatcgttgctacagctactgaggatagatgcctttcccaagcgtggctaacaatctta
 caagaaagactaggaacaggttaagggttgctactatggatcaacttgctgcatgttctggc
 tttatgtattcaatgattactgctaaacaatatatacaatctggtgattacaaacatatt
 30 ttagttgtgggtgctgataaattatctaagattaccgatatgactgaccgttctactgct
 gtattatgtggagacgggtgctggaagctgttgctcatgggagaagttgctgaaggctcgtggt
 atcattagctatgaaatgggttcagacggtagtggtggtaaatacttgctacttagataga
 gaaactggcaaaactcaaatgaatggtagagaagattttaatttgctgtgagaattatg
 ggtgatgcactacgcgtgtagttgagaaagctggtttatcgtctgaagacatagactta
 35 tttgttcacatcaggctaatattagaattatggaatctgcgagagagagattaggaata
 gaaagagaaaaaatgagtgctcagtaataaataatggtaatacttcggctgcctcaata
 ccattaagtattatcaagaattgcaaatggaaaaatcaaagatgacgatacttttagtc
 ttagttggcttcgggtggaggtctaacttggggcgcaatcggtattaaatggggaaaatag

Sequence 3010

MIIVATATGDMFFPSVANILQERLGTGKVATMDQLAACSGFMYSMITAKQYIQSGDYKHI
 LVVGADKLSKITDMTRSTAVLFGDGAGAVVMGEVAEGRGIISYEMSGDGGGKYLILDR
 ETGKLMNGREVFKFAVRIMGDASTRVVEKAGLSEIDILFVPHQANIRIMESARERLGI
 40 EREKMSVSVNKYGNTSAASIPLSINQELQNGKIKDDDTLVLVGFGGGLTWGAIVIKWGK*

45

Sequence 3011

Contig_0564_pos_6613_0
 is similar to (with p-value 9.0e-47)
 50 >sp:sp|P30790|YHI3_RHOCA HYPOTHETICAL 33.7 KD PROTEIN IN HI
 MA 5' REGION (ORF3). >pir:pir|C41608|C41608 hypothetical prot
 ein 3 (himA 5' region) - Rhodobacter capsulatus >gp:gp|M8403
 0|RCAHIMA_3 Rhodobacter capsulatus integration host factor (himA)
 gene, complete cds. NID: g151940.
 55 atgaataaaaaataatagattgttataacgggtatcggaaccttatctccaattggtaac
 gatgctaaaaacaacatgggacaatgcactaaaagggtgttaacgggtatagataaaatcaca
 agaatagatactgatgattataatgtacatcttgctggtgaattgaaagattttaatata
 gaagaccacattgatagaaaagaagctcgccgtatggatcggtttacacaatacgcggtg
 gttgctgcaagagaagcggttaaaagatgcacaattaaatattaatgaaaaaatgcacgac

cgtattggtgatggattggttctggtatcggtggtatggaaactttcgaagttgcacat
 acaacactttagaagaggaccacgtcgagtaagtcatttttcgttccaatgtaatt
 cctgatatggctactggtcaagtttctattgatttaggtgccaaagggcccaatggttct
 acagtaacagcttgtgctacggggactaactcaataggtgaggcattttaaattattcaa
 5 cgtggtgatgcagatgcaatggtgactggtggaacagaggcacctattacacatatggca
 atcgagggttttagtgcaagtcgtgcattatctacaaacaatgacctgaaacagcttgt
 cgaccattccaagaaggccgtgatggcttggttatgggtgaaggggcaggtattgttgta
 cttgaatcattagattcagctaaagagagaggcgctgaaatttacgctgaagttgtaggt
 tatggttccctctggcgatgcacatcatattacagcacctgcgcctgaaggtgaaggtggc
 10 tcacgagctatgcaagctgcttttagatgatgctggaatcaaagctcaagatgtacagtat
 ttaaatgcacatggcacaagtacacctggtggagatttatatgaggttcaagcgattaaa
 aatacattcggtgatgctgcgaagtcattaaaagtaagttcaactaaatcaatgaactgga
 ctttattagggtgctacaggtggaattgaagctatttttctgcgctatcaattcgtgat
 tcaaaggtagccccctacaatacatgcaatcacaccagacgaagaatgtgatttg

15

Sequence 3012

MNKNRNVITGIGALSPIGNDAKTWTDNALKGVNGIDKITRIDTDDYNVHLAGELKDFNI
 EDHIDRKEARRMDFRFTQYAVVAAREAVKDAQLNINEKNADRIGVWIGSGIGGMETFEVAH
 TTLVERGPRRVSPFFVPMILPDMATGQVSI DLGAKGPNGSTVTACATGINSIGAEFKIIQ
 20 RGDADAMVTGCTEAPITHMAIAGFSASRALSTNNNDPETACRPFQEGRDGFMVMEGAGIVV
 LESLDSAKERGAIEIYAEVVGYS SSGDAHHTAPAPEGEGGSRAMQAALDDAGIKAQDVQY
 LNAHGTSTPVGDLIEVQAIKNTFGDAKSLKVSSTKSMTGHLGATGGIEAIFALSIRD
 SKVAPTIHAITPDEECDL

25 Sequence 3013

Contig_0564_pos_1547_414

>pir:pir|I41060|I41060 3-oxoacyl-[acyl-carrier-protein] syn
 thase (EC 2.3.1.41) - Escherichia coli >pir:pir|I84544|I8454
 4 beta-ketoacyl-acyl carrier protein synthase II - Escherich
 30 ia coli >gp:gp|U20767|ECU20767_1 Escherichia coli beta-ketoa
 cyl-acyl carrier protein synthase II (fabF) gene, complete c
 ds. NID: g664869. >gp:gp|AE000210|AE000210_5 Escherichia col
 i K-12 MG1655 section 100 of 400 of the complete genome. NID
 : g1787332. >gp:gp|Z34979|ECFABJ_1 E.coli fabJ gene encoding
 35 beta ketoacyl-acyl carrier protein synthase. NID: g510831.
 gtgatacatcgacaaacaaatcaattaaacactgataacttagaaaaagaagcaacgacaa
 tataaatatgcttttaatttagctgaaattgattctgaaagttttcctatgcacattttt
 agaaaaatagcgaaagacggttttgaagaccatcagctatcactattacaacgaggcgag
 cgtcaagggggaatatattttaaggcaacaaatttcacattacttatttaatagtcgtggc
 40 gtcacttggcaccacaaatcaaattattgttgatcatcaacaagccagttactcgatatg
 ataaccaatttactaaaaaaagaagaatttattattgaacagccaagttatccacctatt
 aaacatacgcttgataaaaaaagggtataagttatattcaagtcagggtgaacaaaatgga
 atacaaatcgacctatttttaatacaataaacaatattttgtatataacaccatctcat
 caatttccaactgggttatgtcaccaatttaaaaaaaagaacacaattaatcaattggtcc
 45 catcaagctaagcaagatatattatcgaagatgattatgattcagaatttagatatattt
 ggcaaacccatacctgcattacaaagtttagacacaaaaggaaaagtcatttatattagt
 actttctcaaaatctttatatccaagctgtaggattgcatatattgttttgccacaaaat
 ttaatgcacaaatataataatcaaaaatataaagagggaataacagtcctgtgcatggt
 caacacatgattgctcaattcatgataagtgaggaaatttgaaagacatttgaataaaatg
 50 cgaaagatatatagagataaacttgatttatattttaaaacgattaaagccctacaatact
 caaa'aaagattgaaggcgactaactggaatgcattttacaataactgttaataatgga
 ttgtcaatgaacaatgtttaaaaaatgcgaaaaaaaataattttaaaattaaaaccttat
 cattacgaaaattattctaaagttatccaaaatttatttttaggatttggggggataaaa
 aaagaagaattagaagatcatgttaatgcattaattcattcactcgttatataa

55

Sequence 3014

VIHRQTNQLNTDNLEKKQRQYKYAFNLAEIDSESFPMHIFRKYAKDVFEDHQLSLLQRGE
 RQGEYILRQQISHYLFNSRGVTCHPNQIIVGSSTSQLLDMITNLLKKEEFIEQPSYPPI
 KHTLDKKGISYIQVPVEQNGIQIDPILNTNNNIIYITPSHQFPTGYVTNLKKRTQLINWS

HQAKQRYIIEDDYDSEFRYFGKPIPALQSLDTKGKVIYISTFSKSLYPSCRIAYIVLPQN
 LMHKYNNQKYKEGNTVPVHVQHMAQFMISGKFERHLNKMRIYRDKLDYILKRLKPYNT
 QIKIEGALTGMHFTITVNNGLSMKQCLKNAKKNNLKLKPYHYENYSKVYPKIFLGFGGIK
 KEELEDHVNALIHSLVI*

5

Sequence 3015

Contig_0566_pos_9236_0

is similar to (with p-value 5.0e-35)

>sp:sp|P94378|3MGH_BACSU PUTATIVE 3-METHYLADENINE DNA GLYCO
 10 SYLASE (EC 3.2.2.-). >gp:gp|Z99123|BSUB0020_157 Bacillus sub
 tilis complete genome (section 20 of 21): from 3798401 to 40
 10550. NID: g2636240. >gp:gp|D83026|D83026_59 Bacillus subti
 lis genome sequence covering lic-cel region. NID: g1783231.
 atgaatatagttgatgtaaataatgattgagaaaatcgacttagcaatagatggcgctgat
 15 gaagtagacagtgcgcttaaccttattaaaggcggtggtggagccttatttagggaaaag
 gtcataagatgaaatggctgaccgatttgcgttgcgttgcgttagatgaaagtaaacctcgtaac
 tatttaggagaaacatttgcattaccagttgaagtcgataaatttaattggtaccaagtt
 gccaaaaaattgagcgtacttatgatattcatgtaagcagaagagttaatgaagatgta
 ccggtttataaccgacaatggttaattacatattagattgttcattgcaaaaatagaattcct
 20 gcttatgagctac

Sequence 3016

MNIVFVNDVEKIDLAIDGADEVDSALNLIKGGGGALFREKVIDEMADRFVVVVDESKLVN
 YLGETFALPVEVDKFNWYQVAKKIERTYDIHVSRRVNEDVPFITDNGNYILDCLQNRI
 25 AYELX

Sequence 3017

Contig_0566_pos_8875_8186

is similar to (with p-value 3.0e-19)

>sp:sp|P32157|YIIM_ECOLI HYPOTHETICAL 26.6 KD PROTEIN IN KD
 30 GT-CPXA INTERGENIC REGION (O234). >pir:pir|S40854|S40854 hyp
 othetical protein o234 - Escherichia coli >gp:gp|L19201|ECOU
 W87_53 E. coli chromosomal region from 87.2 to 89.2 minutes.
 NID: g304961. >gp:gp|AE000466|AE000466_1 Escherichia coli K
 35 -12 MG1655 section 356 of 400 of the complete genome. NID: g
 2367328.
 gtgatatacatgattaaagtgaatgccatatctattggcaaaatagaacattgtcttat
 ggaaactataaaacaaatgcaatcagcggttaaacaatttccttttaaagggtcaaattgtg
 ctcaatcgacttgggttcgtggacgatgaacaagcctatcataaccatggtggtatacat
 40 aaagcgatattgttttagtaaatctaattatcaattatttaaagatgacttagatcaa
 ttacctgaatttgcaatgtttggagagaatttgacagttgaacatctagatgaagcagat
 gtttatttttggttaatcagtatcaactaggcgatacaatcatagaagtatcagatatcgga
 gaaccttgttggaattcaagctaaatatgcaatacctaatttagttcaaaaaa:gtcg
 caatctggtaaaactggattttatttttagagttataaaagaaggatatgtacatcagagt
 45 gataattttaaactcattaaaaaggcagaatcaaacacacgctctatctgtgaaagactta
 aatcatctattctataatgagcgaaataatttaagattaatctatcatgcacttcgaaat
 ccttatctttcacctgatcgaaagaaaaactacagaaaatgaaaacgcgtgccgaaat
 agaaaattcattaaatctgacgataaatag

Sequence 3018

VIYMIKVNAISIGKIETLSYGNYPKMSALNKPFGKQMWLNRLGFVDDQAYHNHGGIH
 KAICCFKSNSYQLFKDDLDQLPEFAMFGENLTVEHLDEADVYFGNQYQLGDTIIEVSDIR
 EPCWKIQAKYAIPLVQKMSQSGKTGFYFRVIKEGYVHQSDNLKLIKKAESNTRLSVKDL
 50 NHLFYNERNNLRIYHALRNPYLSPPDRKKKLQKMKTRAENRKFIKSDDK*

55

Sequence 3019

Contig_0566_pos_3563_2946

is similar to (with p-value 8.0e-19)

>sp:sp|P47968|RPIA_MOUSE RIBOSE 5-PHOSPHATE ISOMERASE (EC 5

.3.1.6) (PHOSPHORIBOISOMERASE)..>gp:gp|L35034|MUSRPI_1 Mus musculus ribose 5-phosphate isomerase (RPI) mRNA exons 1-6, complete cds. NID: g836673.

gtgatataatttggactttataaatcagcagacaacccaactgcaaaagctttattaggt
 5 gttaaaattatttaccagatgactatcaaacatatactggatatattgtagaaactgaa
 gcttatttaggtatacaagataaagctgcacatggttttgggtggcaaaataacacccaaaa
 gtgacttctttatataaaaaaggtggcagcatatgcacatgtgatgcatacgactta
 ttaatcaatttgttacacggactgagggcataaccagaaggtgtacttattcgtgctatt
 gaaccagatgaaggtatcggcgctatgaacgtcaatcgtggaaaatctggatacgagctc
 10 actaatggtccaggaaagtggactaaagctttcaatattccacgatcaattgatggctca
 accttaaatgactgcaattatctatagataccaatcatcgcaaatatccaaaaactatt
 atagaaagtggctgctatcggtattcctaataaaggagaatggacaaaataaccactacgt
 ttcactgttaaaggcaatccatattgtctctagaatgcgcaaatcagattttcaaaatccc
 gacgatacatggaaataa

15

Sequence 3020

VIYLDFFINQTTQTAKALLGVKIIYQDDYQTYTGYIVETEAYLGIQDKAAHGFGGKITPK
 VTSLYKKGGTIYAHVMHLLINLVTRTEGIPEGLVIRAIEPDEGIGAMNVNRGKSGYEL
 20 TNGPGKWTAFNIPRSIDGSTLNDCKLSIDTNHRKYPKTIIESGRIGIPNKGWNTNKPLR
 FTVKGNPYVSRMRKSDFQNPDDTWK*

Sequence 3021

Contig_0568_pos_4867_2714

is similar to (with p-value 3.0e-54)

25 >sp:sp|P45112|RECJ_HAEIN SINGLE-STRANDED-DNA-SPECIFIC EXONU
 CLEASE RECJ (EC 3.1.-.-). >pir:pir|F64110|F64110 single-stran
 ded-DNA-specific exonuclease (recJ) homolog - Haemophilus i
 nfluenzae (strain Rd KW20) >gp:gp|U32801|U32801_1 Haemophilu
 s influenzae Rd section 116 of 163 of the complete genome. N
 30 ID: g1574143.

gtggataatcaggagattcaaaatctatttgaaggcactaacattagtcatttatatg
 ttattaagtgatatgcaaaaagccattgatcgtattaaattagctatcgatcaaaatgaa
 cgaatactagtatatggtgactatgatgcagatggtgttacatctactacaattctagt
 35 tctactttacgtctacttggcgctcagggtgggtggtatattcccaatagatttacagaa
 ggatatggacctaataatgaattagcatttaaaaatgcttatgacgaaggatttccttaata
 ataaactgtagataatggtatacaggacatcatgaaataagtagcatacaagaattaggt
 gtagatgttatagtgacagatcatcatgaaataggagaaaactttacctgatgcttttgca
 attgtacatccgatgcacctaattttgaatatccttttaaatatttatgtggtgcgggt
 gttgcttataaattggcccaaggattgatagagcatccacctcaacatttcatagcatt
 40 gctgccataggtacaattgcagatttagtatcattgacagatgagaatagatatattgta
 aagcaaggagtaaagatatataacaatcatcacccatcgctccataaaggctatcttaaat
 caagcgggttttaatatgatgaaataacagaagaacaattggttttattattggacctcg
 ttaaatgcgggttggttagactagaagatgcacatcattagctgctgaacttttattgtccg
 gaatttgaaggaggcggaatttttagctgaacaagttgaacattttaatcatgaacgtaaa
 45 gatatagtatctaaaattactaatgaagcattgttattagcagagggaacaaatcaagcaa
 ggccatttgtttcttttacttgccaaagaggggtggcatgagggtgtatttaggtattgtt
 gcacataaaattgtagaaacatatgcactacctacattaatttttaaatattgatgaaat
 caaaatcatgccaaagggttctgcgaggtcgattgaacaagtttccatggttgatatttta
 aatgatcatcaacagtttaattgataagtttgggtggtcatcacatggctgcaggaatga
 50 atgtctatcgataatattgaacatttacataaaagagctagatatgtggatgaaagaacta
 actgttaccacttcatttagagccttcaataaagggtggatgcacaactgaagaaaaagaa
 attaacattaaaaatattaaagatatatttcaattaaggccttttggtacggactttaat
 agtcctctttttatggttagagatctaattgtcaagtcaacaaagggaattggacaggat
 aataagcatcttaagttaacgcttggtcattcaggtttaactgctttattttggaatcat
 55 ggacatttggcaagtgaacttgaaccagggtcaaccgattcatataataggaacattgc
 aatgaatgaatggaatggaatcaaacacctcaatttatcatcaaagacattgctatagac
 caattacaatttttagattatcgtagtaaacgtaaaaatatacaatttaagaaactgaa
 tcaaatgttgctatgtcattcatccgaaacttaaaaaagcaattcacattactatcat
 tatggtgagggttattgatagaccttatgataaaatagtgttttagagatttacctaact

atgggttgaattgaacaaaccttagaacattcacaaatcttcaactatatttagttctg
 cagcatgaaaagtcaatatattttgaaggtatccctagcaagtcgctttttaaagtgt
 tacaaagctttaataaacaacaaaagaactgacttaattaaagaaggtatgctgctttgt
 gagtactttaatatattaagccagaaatattaacattcatgcttaaagtctttaagaattg
 5 gagtttatttacgacgaaaagggttgattaagataaatccagcgccaaataaacaagat
 atagaaaatagtcgtattttaccaaatgagacgagcacgtatggaagtagaagaacgtctt
 ctctatgatgatttttttaatatataaagaatggataatatcaaagttaacatag

Sequence 3022

10 VDNQEIQNLFEGTNISHDYMLLSDMQKAIDRIKLAIQNERILVYGDYDADGVTSTTILV
 STLRLLAGQVGWYIPNRFTEGYGPNELAFKNAYDEGISLIITVDNGIQGHHEISTIQLG
 VDVIVTDHHEIGETLPDAFAIVHPMHPNFYFPFKYLCGAGVAYKLAQGLIEHPPQHFIAL
 AAIGTIADLVSLTDENRYIVKQGVKILNNHTPSSIKAILNQAGFNDEITEETIGFIIGPR
 LNAVGRLEDASLAAELLSDEFEEAEFLAEQVEHPNHERKDIVSKITNEALLLAEEQIKQ
 15 GHLFLLLAKEGWHEGVLGIVASKIVETYALPTLILNIDENQNHAKGSARSIEQVSMFDIL
 NDHQQLIDKFGGHHMAAGMTMSIDNIEHLHKELDMWMKELTVTTSLEPSIKVDAQLEEKE
 INIKNIKIDIFQLRPFGTDFNSPLFMVRDLIVKSTKGIGQDNKHLKLTGLHSGLTALFWNH
 GHLASELEPGQPIHIIGTLQINEWNGNQTPQFIKDIAIDQLQILDYRSKRKNIQFKETE
 SNVAYVIHPKLKKSNSHYHYGEVIDRYPYDKIVFRDLPTMVEIEQTLEHSQISQLYLVL
 20 QHEKSIYFEGIPSKSLFKKCYKALINKKETDLIKEGMLLCEYLNKPEILTFMLKVFKEL
 EFIYDEKGLIKINPAPNKQDIENSRIYQMRARMEVEERLLYDDFLNIKEWIISKLT*

Sequence 3023

Contig_0569_pos_4529_5401

25 is similar to (with p-value 3.0e-33)
 >sp:sp|P39312|CYCA_ECOLI D-SERINE/D-ALANINE/GLYCINE TRANSPORTER. >pir:pir|S56433|S56433 hypothetical protein o470 - Escherichia coli >gp:gp|U14003|ECOUW93_120 Escherichia coli K-12 chromosomal region from 92.8 to 00.1 minutes. NID: g126317
 30 2. >gp:gp|AE000492|AE000492_4 Escherichia coli K-12 MG1655 section 382 of 400 of the complete genome. NID: g1790649.
 atgcctgaactacctgaagttgaacatggttaaaagaggtattgagccatttataaaaagt
 gcaaaaatagagaaagtaacttttgctaaaaatgtaattaacggtagaataataaccgt
 gagacgattataaaaggtatggaattagatacttttaaaaaacttactgaagggtatggt
 35 ataaaaaaagttgaagaagaagtaagtaacattttttatatagcggatcatgacgat
 gatagaatccttagttagtcatttaggtatggcaggcggattccttgttgaataaacct
 gatgagataagtaacacgaattatcgaaagcattggcaagtcattttcgatttgataat
 aaacaaaaatagtcatttctgatatacagacggtttggagaaattagaaatatagtcatt
 tttgatagttatccatctttattagaaatcgctccagaaccatttgaagaggtggcattt
 40 gaacactatttagaatgtttgacaatgaaaaaatataagaataaaccataaaacaaacg
 attccttgatcatcggttatagcaggagctggaaatatctatgcctgtgaagctttatc
 agagctggtattactccggataaaattactaattcactcactaaacaagaaagaaaatcc
 ctcttttattatgttcgagaagttttagaagagggtataaaatatggaggtagtagtatt
 cgagattataggtatcgagatggtaaaactggacaattgcaattacatttaaatgtatat
 45 aaacaaaaaaagtgcaaggtttgtggtcattcgattgaaacaaaagtgatagctggtaga
 aatagtcatttttgccaaactgtcagagataa

Sequence 3024

MPPELPEVHVKGRIEPIKSAKIEKVTFAKNVINGKNNNRETIKGMELDTFKKLTEGYV
 50 IKKVERRSKYIIFYIADHDDRIIVSHLGMAGGFVNNLDEISTPNYRKHWQVIFDLN
 KQKLVSIDIRRFGEIRNIVNFDSPSLLEIAPEPFEEVAFEHYLECLTMKKYKNKPIKQT
 ILDHRVIAGAGNIYACEALFRAGITPDKITNSLTQERKSLFYVREVLEEGIKYGGTISI
 SDYRHADGKTGQMQLHLNVYKQKKCKVCGHSIETKVIAGRNSHFCPCNQR*

Sequence 3025

Contig_0569_pos_8390_8887

is similar to (with p-value 1.0e-38)
 >sp:sp|P44948|FPG_HAEIN FORMAMIDOPYRIMIDINE-DNA GLYCOSYLASE (EC 3.2.2.23) (FAPY-DNA GLYCOSYLASE). >pir:pir|A64104|A6410

- 4 formamidopyrimidine-DNA glycosylase (fpg) homolog - Haemophilus influenzae (strain Rd KW20) >gp:gp|U32776|U32776_1 Haemophilus influenzae Rd section 91 of 163 of the complete genome. NID: g1573969.
- 5 gtgattaaacttgggagggtgacatatatgaaatgcccaaatgtaattctacacattcc
agagtgggttattcaagacatgcagatgaggccaatgcgattagacgtagaagagaatgt
gaaaattgcggaacgcgttttacaacatttgaacatattgaagttagtccattaatagta
gtgaagaaagatgggactagagaacaatttttaagagaaaaatattaaatgggttagta
agatcttgcgagaaacgaccagtacgttatcaacaacttgaagacataactaataaagt
10 gagggtgcaacttagagatgagggacaaactgaaatttcttctagagaaattggagagcat
gttatgaatttgttaatgcattgtgaccaagtttctctatgtaagatttgcattctgtatat
aaagaattcaagatgttgatcaactcttagagtcaatgcaaggatcttgagtataat
aaacggagtataaatag
- 15 Sequence 3026
VIKLGRVITYMKPCKNSTHSRVVDSRHADEANAIIRRRRECENGTRFTTFEHIEVSPLIV
VKKDGTRQFLREKILNGLVRSCEKRPVRYQQLLEDITNKVEWQLRDEGQTEISSREIGE
HVMNLLMHVDQVSIVRFASVYKEFKDQVQLLESMQGILSDNKRSDK*
- 20 Sequence 3027
Contig_0569_pos_1314_691
is similar to (with p-value 1.0e-33)
>sp:sp|P46227|YRS1_SYN6 HYPOTHETICAL 19.1 KD PROTEIN IN PS
BB-RPS1 INTERGENIC REGION (ORF 168). >pir:pir|S51484|S51484
25 hypothetical protein 168 - Synechococcus sp. >gp:gp|D28752|S
YORPS1_2 Synechococcus sp. gene for ribosomal protein S1, complete cds. NID: g560122.
atgtctgttgtaccatggcaacaattgaatcctgctgacagtcacattgtaaaatgttt
ggattagttggaatcccttttgcagcaggtattattaactttgttgtacttacagctgca
30 gcctcttcttgaatagtggtatatttgcataagccgtacgatgtttggattagctgga
agaaagcaagggtccagcattcttcatagacaaccaataagcagcggtaccacattatgct
attttagtgacatgtggcttattaagattttcagtcgtgttaaatgcaatttttaagat
gcgactaaagtgttcgtacaaattacaacattttcaactgttttaaatattatgatttgg
acaattattatgatcgctatctaggttatttaagacatgaaccgaaacagcataaagaa
35 agtaactataaaatgtggggcggaatacatggcttacagttatttaggggtcttttgc
tttatttttattatactattgattaatagtgcaacgcgttatgccgtactttctgcaccc
gtatgggtttgtatcatgctattgatgtatcaaaaataaaaaagaatctcgcaagct
aaaattaaaaatgaggaagagtaa
- 40 Sequence 3028
MSVVPWQQLNPADSPYVKMFGLVGIPFAAGIINFVLTAAASSCNSGIFANSRTMFGLAG
RKQGAFLHRTNKHGVPHYAILVTCGLLSISVVLNAIFKDATKVFVQITTFSTVLNIMI
TIIMIAYLGLYLRHEPKQHKEKNYKMWGGKYMAYSILGFFAFIFIILLINSATRYAVLSAP
VWFVIMLLMYQYKKESRKAKIKNEEE*
- 45 Sequence 3029
Contig_0570_pos_0_488
is similar to (with p-value 5.0e-39)
>sp:sp|P54715|PTIB_BACSU PTS SYSTEM, ARBUTIN-LIKE IIBC COMPONENT
50 ONENT (PHOSPHOTRANSFERASE ENZYME II, BC COMPONENT) (EC 2.7.1
.69). >gp:gp|Z99108|BSUB0005_89 Bacillus subtilis complete genome
(section 5 of 21): from 802821 to 1011250. NID: g2633055. >gp:gp|D50543|D50543_3 Bacillus subtilis DNA for 76-deg
ee region, complete cds. NID: g1486240.
55 atgtttgctttttcggtattgtttgggattcgctacattatttataaaatccaaccatt
atgggaggatttagctgatcagcaaacattttggtttaaattttggtctgttattgaatca
gggtggtgggtaataatttacacatatggaaattgtctttgtagttggcttaccattatct
cttgctaaaaaagcaccaggacatgcagcttttagcagctctaattgggatatttaattgtt
aatacttttatcaatgcaattttaactcaatggccacatacttttggcgctaatttataaa

aaaggtgtagaaaacacaacaggattaaaatcgattgcaggtattgaaacgtagatacc
aatatttttaggtgcaatcattatctcaggaataataacgtggatacataatagatattac
agtaagcgtttacctgaaatgttaggtgtatttcaaggattaacattcgttgtaacaatc
tctttctt

5

Sequence 3030

MFAFFGIVLGFATLKFKNPTIMGGLADQQTFWFKFWSVIESGGWVIFTHMEIVFVVLPLS
LAKKAPGHAALAALMGYLMFNTFINAILTQWPHTFGANLKKGVENTTGLKSIAGIETLDT
NILGAIISGIITWIHNRYYSKRLPEMLGVFQGLTFVVTISFX

10

Sequence 3031

Contig_0580_pos_966_1382

is similar to (with p-value 5.0e-85)

>sp:sp|P43848|PUR5_HAEIN PHOSPHORIBOSYLFORMYLGLYCINAMIDINE
CYCLO-LIGASE (EC 6.3.3.1) (AIRS) (PHOSPHORIBOSYL-AMINOIMIDAZ
OLE SYNTHETASE) (AIR SYNTHASE). >pir:pir|G64122|G64122 5'-ph
osphoribosyl-5-aminoimidazole synthetase (purM) homolog - Ha
emophilus influenzae (strain Rd KW20) >gp:gp|U32822|U32822_2
Haemophilus influenzae Rd section 137 of 163 of the complet
e genome. NID: g1574265.

20

atggaagcacaaatggaaaaagatggtaattactatatggaaggaatattagatgatatt
caacaagatggatattggtttcttaagaaccgttaactattctaaagtgagaaggatatt
tatatttctgcaagccaaattcgacgttttgaaataaaacgtggtgataaagtaacgggt
aaagttcgtaaaaccaaagataatgaaaaatattatggcttacttcaagttgattttgta
aacgaccataatgcagaagaagtaaaaaacgtccttcacttccaagctttaacacctct
ttatccggaagaagaatcctattagaacgcaatctacaaattattccactcgtattat
ggatttagtcacaccaataggtcttggtcaacgtgggtcttatagttgcaccacctaa

25

Sequence 3032

MEAQMEKDGNYMEGILDDIQDGYGFLRTVNYSKGEKDIYISASQIRRFEIKRGDKVTG
KVRKPKDNEKYYGLLQVDFVNDHNAEEVKKRPSLPSFNTSLSGRKNPIRNAIYKLFHSYY
GFSHTNRSWSTWSYSCTT*

30

Sequence 3033

Contig_0580_pos_2122_3000

is similar to (with p-value 1.0e-27)

>sp:sp|P38527|RHO_THEMA TRANSCRIPTION TERMINATION FACTOR RH
O. >gp:gp|L27279|TMORHO_1 Thermotoga maritima rho gene, comp
lete cds. NID: g454858.

40

atgaaagcgccaggttctggtatcaggtactgacggtgtgggtacaaaagttaaaattagca
attgactatggaaagcatgacacaattgggtattgatgctgtcgcaatgtgtgtaaattgat
attttaacaacaggtgctgaacctttatacttttttagactatattgccacgaataaagta
gtgccaaagtactatagagcaaatcgttaaaggtataagtgacggttgcaacaaaccaat
acggcacttataggcgggtgaaactgctgaaatgggagaaatgtatcatgaaggtgaatat
gatattgctggttttgcagtaggagcggtagagaaagaggactatattgatggttcaaat
gttgaagaaggacaagcaattattgggttttagcttcaagtggtattcattcaaatggctat
agtctagttagaaaaatgataaaagaatcaggagttcaattacatgatcaatttaattggt
caaacctttttagaaaccttccttgccacacaaaattgtatgtaaagcctattcttgaa
ttaaagaaacatattgatatcaaagcgatgagccatattactgggtggaggtttctatgaa
aatattccgcgtgcccttcctaaaggtttatcagcaaaaatagatacacaaatcattccca
acgttggaagtctttaattggcttcaaaaaacagggaacatttcaacgaatgaaatgtat
aacatatttaatatgggtattggatatacaattattgttgacaaaaaagatgttcaaaca
acattaacaacggttacgtgcaatggatacaactgcatatgaaattgggtgagattataaaa
gatgatgatacacctattcattttattggaggtagaatag

55

Sequence 3034

MKAPVLVSGTDGVTGKLLKLAIDYGKHDITIGIDAVAMCVNDILTTGAEPFLDYIATNKV
VPSTIEQIVKGISDGCEQNTALIGGETAEMGEMYHEGEYDIAGFAVGAVEKEDYIDGSN
VEEGQAIIGLASSGIHSNGYSLVRKMIKESGVQLHDQFNGQTFLETFLAPTCLYVKPILE

LKKHIDIKAMSHITGGGFYENIPRALPKGLSAKIDTQSFPTLEVFNWLQKQGNISTNEMY
NIFNMGIGYTIIVDKKDVQTTLTTLRAMDTTAYEIGEIKDDDTPIHLLVE*

Sequence 3035

5 Contig_0581_pos_6447_3826
is similar to (with p-value 4.0e-19)
>gp:gp|X99710|LLVSPFEP_4 L.lactis ORF, genes homologous to
vsf-1 and pepF2 and gene encoding protein homologous to met
hyltransferase. NID: g1771200.

10 atgaaagctagtgaatttaggcaaaaatatttgaatttctttgtagaaaaaggacatatg
attgaaccgtctgcaccacttgtacctatcgatgatgattcattattgtggatcaattct
ggtgtagctaccttgaaaaatactttgacggacgcgaaactcctaaaaaccaagaatt
gtcaattctcaaaaagctatacgaacaaacgatatagaaaacgtcggttttactgctcgc
catcatacttttttgagatgctaggtaatttttcaatcggtgactactttaaacatgaa
15 gcgattgaatttgcctgggaatttctaacaagtataaatggatgggcatggaacctgag
aagctatatgtgactattcatcctgaagatactgaagcttttagaatttggcatgaagac
attggttttagaggaaagtcgcatcattcgaatagaaggcaatttttgggatatcggggaa
ggaccatctggaccaaatacagaaaatttctatgatcggtggatcggcttatggaaaagat
gatcctgctgaagaaatgtatcctgggtgggtaaaatgaagatatttagaagtttggat
20 ctagtgttttagtgagtttaatacacaataaagacaatacttacacaccactaccaaataaa
aatattgatactggcatgggattagaacgtatgacgtctatctcacaataatgtaagaaca
aattatgaacagacttatttatgcctataattaaggaagtagaacatgtttcaggaaaa
aaataatttaattgatgatgcacaagatgttgcatttaagttatcgcagaccacattaga
acgatttctttcgcaattgctgatggcgcatcaccagctaataagggttagaggatatgta
25 ttaagaagattattacgcagagcagttcgctttagccaatcattaggaattaatgaacca
tttatgtataaaacttggtgatatagtcgctgatcatggaaccatattatccaaatgtc
aaagacaaatccaactttattaaacgtgtcattaaatcagaggaagaacgcttccatgaa
acgcttgaggaaaggtcttacgattttaaatgaactgataaaagaagcgaagaatagtgat
caggttatttaaaggtcatgatgcttttaagttatatgatacttatggattcccaatagaa
30 ttaactgaagaatttagcaactcaagaaaatttgcctgttgatagcctactttgaacag
gaaatgcaacaacagagagatcgagctagacaagctcgacagaattctcaatcaatgcaa
gttcaaaagtgaagtactaaaaaatattcaagatgaagtcattttgttggctatgaaact
acggactatcaatcattaataactcatatcatatacaatgggtgaagaagttaaacatgtt
gaagcaggagaacaatttactttattttaagagaacgcctttctatcgagtaagtggt
35 ggacaggctcgagataagggaacagttggtaatgagagctttgaaataaatgtaactgac
gtaactaaagcgctaatggccaaaacttacacaaaggtattgtgcaatttggatgaagca
acacagaacgcgaaagtagaagcacgtgttaacaaagaggatagacgacttattcaaaaa
aatcatagtgtcacacatttattacatgctgcattaaaagaagtattaggagatcatgtt
aatcaggctggttcggttagtagaacctgaaagactacgttttgatttctcacattttggt
40 cctatgacacagaagaataattatttagtagaacgtagagtaaatgaagaaatttgagaga
gctatcgacgtccgtattcaagaaaatgagttatgaagaagccaaatcaataggcgctatg
gctttatttgggtgaaaaatatggagatattgttcgctgtgcaatatggcaccattttca
atagaattgtcggtggaatacatgtggaataataactcggaatttggctctcttaagatt
gtgagtgaatctggaacaggtgcccgtgttagaagaattgaagctttaacagggtaaaggt
45 gcattcttacatcttgaagaaattgaaacacagtttaataatattaaaaatcatttaaaa
gttaaatccgataaccaagtagttgaaaaagttaaacaacttcaagaagaagaaaaagga
ctgcttaacaattagaacaacgcaacaagaaataaatcactaaagatggggaacatt
gaagaacaggttgagttgattaataatttgaaagtttttagcaacagaagtagaaattcca
aatccaaaagcgatacgttcaactatggatgactttaaatctaaacttcaagatactatt
50 atagtgttagtcggacaagttgatggaaaggtttctgtaattgctacagtaccaaataca
cttacaatacaagtaaaaagctggagatcttcaaaaacatgacaccaattatttggtgga
aaaggtggaggtcgctctgatatggctcaaggtggcggaactcaacctgaaaaataaca
gaagcattacgctttattaaagattacattaaaaatctataa

55 Sequence 3036

MKASEIRQKYLNFVEKGHMIEPSAPLVPIDDDSLWINSQVATLKKYFDGRETPKKPRI
VNSQKAIRTNDIENVGFTARHHTFFEMLGNSIGDYFKHEAIEFAWEFLTSDKWMGMEPE
KLYVTIHPEDTEAFRIWHEDIGLEESRIIRIEGNFWDIGEGPSGPNTTEIFYDRGSAYGKD
DPAEEMYPGGENERYLEVWNLVFSEFNHNKDNTYTPLPNKNIDTGMGLERMTSISQNVRT

NYETDLFMPPIKEVEHVSGKKYLIDDAQDVAFKVIADHIRTISFAIADGALPANEGRGYV
 LRRLLRRAVRFSQSLGINEPFMYKLVDIVADIMEPYYPNVKDKSNFIKRVIKSEERFHE
 TLEGLTILNELIKEAKNSDQVIKGHDAFKLYDTYGFPIELTEELATQENLSVDMPTFEQ
 EMQQQRDRARQARQNSQSMQVQSEVLKNIQDESQFVGYETTDYQSLITHIYNGEEVKHV
 5 EAGETIYFILRETPFYAVSGGQVADKGTGTVGNESFEINVTDVTKAPNGQNLHKGIVQFGEA
 TQNAKVEARVNKEDRRLIQKNHSATHLLHAALKEVLGDHVNQAGSLVEPERLRFDSSHFG
 PMTQEEINLVERRVNEEIWRAIDVRIQEMSIEEAKSIGAMALFGEKYGDIVRVVNMAFVS
 IELCGGIHVNNATAEIGLFKIVSESGTGAGVRRIEALTGKGAFHLLEEIETQFNINIKNHLK
 VKSDNQVVEKVKQLQEEEEKGLLKQLEQRNKEITSLKMGNIEEQVELINNKKVLATEVEIIP
 10 NPKAIRSTMDDFKSLQDTIIVLVGQVDGKVSIVATVPKSLTNQVKAGDLIKNMTPIIGG
 KGGGRPDMAQGGGTQPEKITEALRFIKDYIKNL*

Sequence 3037

Contig_0581_pos_3500_3072

15 is similar to (with p-value 5.0e-25)
 >gp:gp|U46071|RCU46071_3 Rhodobacter capsulatus cytochrome
 c biogenesis (cycH) gene, complete cds, and sarcosine oxidas
 e gene, partial cds. NID: g1353871.
 atgctaaagcataaaatatttaggactagatggttggaagtaaaactggtggcattgctata
 20 agtgaccttatgggttgactgctcaagggtagacacactccgtatcaacgaagaacaa
 gatgatttaggaattgatcaactcgtgaagattattaaagataatcaagttggaactggt
 gttattggcttgcccaagaatatgaacaattcaataggttttcggggagaggcttcaata
 aaatataaagaaaagttacaagagtctatcccttctattgatattggtatgtgggacgaa
 cgtttaagtacaatggctgctgaaagatctttacttgaagcagatgtttcaagacaaaaa
 25 agaaaaaaggtaatagataaaatggcagctgtattttattttacaaggctatttagattct
 attcaataa

Sequence 3038

30 MLKKHILGLDVGSKTVGIAISDLMGWTAQGLDTRLINEEQDDLQDLVKIKNQVGTW
 VIGLPKNMNNISGFRGEASIKYKEKLQESIPSIDIVMWERLSTMAERSLLEADVSRQK
 RKKVIDKMAAVFILQGYLDSIQ*

Sequence 3039

Contig_0581_pos_2179_1685

35 >sp:sp|P00957|SYA_ECOLI ALANYL-TRNA SYNTHETASE (EC 6.1.1.7)
 (ALANINE--TRNA LIGASE) (ALARS). >gp:gp|D90892|D90892_9 E.co
 li genomic DNA, Kohara clone #446(60.5-60.9 min.). NID: g180
 0074. >gp:gp|AE000353|AE000353_11 Escherichia coli K-12 MG16
 55 section 243 of 400 of the complete genome. NID: g1789037.
 40 atgaatagaactaaaaatatttttagaaattggtacagccattggttatagttcaatgcaa
 ttcgctaataatttcgaaagatataaatgtaacaacaattgaacgaaatgaagacatgatt
 catcttgcataaaaaagttcataaaaaagtatcgataccagaatcaaatccgtttaattgaa
 tacgatgctttgaatgcatttgaacaagtcaatgacaaacaatatgatatgatattttatc
 45 gacgcagctaaagcaccaatcaatgaaatttttccaactatatacaccgttattaaaaaaa
 ggtggaattgtggttactgataatgttttatatcatggatttgtttcaaataatagacggt
 gttcgttcgagaaatgtgaaacaaatgggttaagaaagtgaacagtcacaaatgaatgggtg
 atggagcaatctcaatttacaacaaactttataaatatggatgatggattagcaatatct
 ataaaaggagaatga

Sequence 3040

50 MNRTKNILEIGTAIGYSSMQFANISKDINVTTIERNEDMIHLAKKFIKKYRYQNQIRLIE
 YDALNAFEQVNDKQYDMIFIDAAKAQSMKFFQLYTPLLKKGGIVVTDNVLYHGFVSNIDV
 VRSRNVKQMVKVQYQYNEWLMEQSQFTTNFINMDDGLAISIKGE*

Sequence 3041

Contig_0584_pos_9628_10956

is similar to (with p-value 2.0e-23)
 >sp:sp|P75822|YBJT_ECOLI HYPOTHETICAL 53.7 KD PROTEIN IN AR

TP-POXB INTERGENIC REGION.

atgatgccatcagcaaaattaacgcaggctagctttgaagatatggatgcattgtagca
 gataattttgcgcgtgcggcgcaaaagcagggtgttaagcatattgtctacatgagtgg
 ttaataccagaaaaatgatgagctatctgcacatttaagaagtcgacttgaatgtgaaaaa
 5 attcttgggtgattacgggtataccagtttagcacattacgtgcagggttaattattggtgca
 aaaggaagttcttatccaattcttaaacgactagttaagagattgccagcaatggttttg
 cctagtgtgggttacaataaaaattgcacctgttgccattgacgatgtgatagatggttta
 gcagcatgtgtgaatcgaaccccaagataacgaagcaattgatatcacaggctcctgaa
 gtgatgaattataaaacgctgatacagcgcacagctaacgtacttgataagcgactgcct
 10 atgcttgatttacctattataccattatcataagtcggtattgggtacaactgatttca
 aatgtaccgaagaatgggtatataccattaatgaatagtttaactcacgatatggtacca
 catcgaacgcggttgtgtctaaactgtccgtaggaaatatcacctttgaagatagtgtg
 aaaagagcactaagagaagaacaaaagacttctaagaaaaagtcggattcgaaaaattct
 caatcatttgggcgtatgcatcaagaaattaaagatgtacgagccattacacggtttaaa
 15 attccggaaggttattcgattaaagatgtgactaaagaatatgcaaaattcatcaataat
 atcacactacatctcgttaaaggtagcataaataaacgagaatttaatatgaatttgccc
 ttcattataaatttttttaaaaatggaacgtgatgaagctgactctacagaagatatg
 gtggtatataaatattgtgggtggcgatttagcacattcaaatgatgggtggaaatgcacgc
 tttgaattccgaagaataagaacaccaaagagggtattattgctttacaagaatatgaa
 20 cctacattaccttgggtagtagtataaaactaactcaagctaaagcacacaagactgttatg
 aatatttttaaaaataaaaatggcacgtttatcgcaacaaaaaaatgtgaaagatgaaaca
 tatatgtctaactcgtgtaactattggagtaacagtagcatcagcgttcgttattggaagc
 gcagtaggggttccaacttttttaaaaagcatcaaatcaaaaagaacacaatgtcgaatgca
 gaattataa
 25

Sequence 3042

MMPSAKLTQASFEDMDALLADNFARAAQKQGVKHIYVMSGIIPENDELSAHLRSRLECEK
 ILGDYGIPIVSTLRAGLIIGAKSSYPILKRLVKRLPAMVLPWAYNKIAPVAIDDDVIDGL
 AALVNRTPKDNEAIDITGPEVMNYKTLIQRTANVLDKRLPMLDLPIPIIISRYWVQLIS
 30 NVPKEMVYPLMNSLTHDMVPHRKRVVSNLSVGNITFEDSVKRALREEQKTSKKKSDSKNS
 QSFGRMHQEI KDVRATIRFKIPEGYSIKDVTKEYAKFINNITLHLVKGTINEREFNMNLP
 FINKFILKMERDEADSTEDMVVYNIVGGDLAHSNDGGNARFEFRIRNTNEGIIALQEYE
 PTLPWVYKLTQAKAHKTMVNI FKNKMARLSQQKNVKDETYMSNRVTIGVTVASAFVIGS
 AVGFQLFKKHQIKKNTMSNAEL*
 35

Sequence 3043

Contig_0585_pos_4328_4747
 is similar to (with p-value 7.0e-17)
 >sp:sp|P77279|YBBL_ECOLI HYPOTHETICAL ABC TRANSPORTER ATP-B
 40 INDING PROTEIN IN USHA-TESA INTERGENIC REGION. >gp:gp|U82664
 |ECU82664_88 Escherichia coli minutes 9 to 11 genomic sequen
 ce. NID: g1773084. >gp:gp|AE000155|AE000155_6 Escherichia co
 li K-12 MG1655 section 45 of 400 of the complete genome. NID
 : g1786692.
 45 atgcaacaaagtgaagtttaacggttatacaattgaagataatatgaaatttctgctgag
 gctagaagtgaagcttttgaccgtgataaagcgaacaaactcatctctcaagtaggatta
 ggtaattatcagtttagatgctcaaattgagcacatgtctgggggagagcaacaacgtatt
 accatcgctagacaactcatgtatgaacctgaagttttattattggacgaagctactagc
 gctttagatacacataataaaaaagaaaaattgaagaaattatatttaaacttagcagataaa
 50 gggattgccattttgtggattacgcatagtgatgaccaaagtatgcgtcatttttaagcgt
 agaatcacaattactgacggtaagatatcgagtgatgaggagttgaatggtaatgagtaa

Sequence 3044

MQQSELIGYTIEDNMKFPAEARSEAFDRDKAKQLISQVGLGNYQLDAQIEHMSGGEQQR
 55 TIARQLMYEPEVLLLDDEATSDTHNKKKIEEII FKLADKGIAILWITHSDQSMRHFKR
 RITITDGKISSDEELNGNE*

Sequence 3045

Contig_0585_pos_5112_5516

is similar to (with p-value 2.0e-32)

>sp:sp|P77307|YBBM_ECOLI HYPOTHETICAL 28.2 KD PROTEIN IN US
HA-TESA INTERGENIC REGION.

5 atgttagctaataatgggttaatcgcgattaatctcgcttatcagaatttagaaagagca
tttgttcaagatgtttctgataattgaatccaaacttacgtagcagcgacacctaagctc
gcatcaaaatcagctattagagaaagtatacgcttagcaattgttctacaattgattct
gtaaaaacatatgggtctagtttcaattccagggtatgatgacaggattgattatcgaggc
gttgaccacttcaagcaattaaatttcaattgcttgctgctgtttattcatacaacagcg
acgattatgtctgcactcattgcaacgtatatgagttacggtcagttctttaatgtctcg
10 catcaactcattgctagaacgcacgcacaagacaaagtagttaa

Sequence 3046

MLANNGLIAINLAYQNLERAFVQDVSDIESKLTAAATPKLASKSAIRESIRLAIVPTIDS
VKTYGLVSI PGM MTGLI IGGVDPLQAIKFQLLVFIHTTATIMSALIATYMSYGQFFNAR
15 HQLIARTQTRQSS*

Sequence 3047

Contig_0589_pos_999_559

is similar to (with p-value 1.0e-35)

20 >gp:gp|Y10549|BSFMS_1 B.stearothermophilus fms gene. NID: g
2266413.

atgataacaatgaaagatattataagagatgggtcatccaacacttcgtgaaaaagcgaaa
gaattaagcttcccactttctaacaatgataaagaacattgcgcgcaatgcgtgaattt
ctaatacatagtcaggatgaagaaacgcgcaaaacggttatgggttacgttctggcgtaggt
25 ttagctgctccacaaattaatgaacaaaacgtatgattgctgtctacttacctgatgat
ggaaacggtaaatcggtatgattatatgctcgtaaatcctaaaataatgagttacagtgt
caagaagcttatttaccactggcggaaggttgcttaagtggtgatgaaaacatcccaggt
ttagtgcatcgctcatcatttcttaatatcactttcggttttaaaaccacaaaatttaact
atcatagagattcttctttaa

30

Sequence 3048

MITMKDIIRDGHPTLREKAKELSFPLSNNDKETLRAMREFLINSQDEETAKRYGLRSGVG
LAAPQINEPKRMIAVYLPDDGNGKSYDYMLVNPKIMSYSVQEAYLPTGEGCLSVDENIPG
35 LVHRHFLISLSVLKPQNLTIIETLL*

35

Sequence 3049

Contig_0591_pos_1336_2091

is similar to (with p-value 7.0e-70)

40 >sp:sp|Q06753|YACO_BACSU HYPOTHETICAL RRNA METHYLASE IN CYS
S 3' REGION. >gp:gp|D26185|BAC180K_157 B. subtilis DNA, 180 k
ilobase region of replication origin. NID: g467326. >gp:gp|Z
99104|BSUB0001_96 Bacillus subtilis complete genome (section
1 of 21): from 1 to 213080. NID: g2632267.

gtgaatgtggaagatatagtgatagtaggtagacacgcagttaaagaagcaattatatca
45 ggtcacgccataaataagattttgattcaagacggtataaaaaagcaacaaattaacgac
attttaaaaaatgcaaaatcacaaaaattaattgtacaaacggtacaaaatctaaatta
gatttttttagcaaatagcacctcaccaggggtgtggctgcttttagtagcccatatgaatat
gcaaacttcgatgaatttttcaaaaaacaaaagaaaaagcccggtattcaactgttatc
attttagatgggttagaagacccgcataatcttggtcttatattaagaacagcagatgct
50 tctgggtgttgatgcggttattatacctaaaagacgatcagttgcgctaacacagaccgt
tcaaaaagcttctacaggagcgattcagcatgttccggttataagggttactaatctttcg
aaaactatcgacgaattaaaagacaacggcttttggttgctgacgaggaagctaataat
gcaacggattatagagatttacaagcagatatgtcactaggtattgtaataggtagtgag
gggcaaggtatgagtcgttttagtgagtgataagtgatgtttcatattaagattccaatg
55 gttggacatgtcaatagcttgaacgcgtctgtggctgcaagtttaatgatgtatgaagta
tatcgtaaacgtcatcagttagaggaaaagtcatga

Sequence 3050

VNVEDIVVGRHAVKEAIIISGHAINKILIQDGIKKQQINDILKNAKSQKLIVQTVPKSKL

DFLANAPHQVAAALVAPYEYANFDEFLQKQKKKARYSTVIILDGLEDPHNLGSILRTADA
SGVDAVYIPKRRSVALTQTVAKASTGAIQHPVIRVTNLSKTIDELKDNGFWIAGTEANN
ATDYRDLQADMSLGIVIGSEGGMSRLVSDKCDPHIKIPMVGHVNSLNASVAASLMMEYEV
YRKRHQLEEK*

5

Sequence 3051

Contig_0592_pos_3596_3057

is similar to (with p-value 3.0e-43)

>sp:sp|Q47155|DINP_ECOLI DNA-DAMAGE-INDUCIBLE PROTEIN P. >g
10 p:gp|D38582|ECODINJ_11 Escherichia coli genes for 'YafH, Yaf
I, YafJ, YafK, YafQ, DinJ, YafL, YafM, FhiA, MbhA, DinP, Yaf
N, YafO and YafP. NID: g984576. >gp:gp|D83536|ECOTSF_39 Esch
erichia coli genome, 4.0 - 6.0 min region. NID: g1208942. >g
p:gp|U70214|ECU70214_72 Escherichia coli chromosome minutes
15 4-6. NID: g1552727. >gp:gp|AE000131|AE000131_10 Escherichia
coli K-12 MG1655 section 21 of 400 of the complete genome. N
ID: g1786415.

atggattatttttttctcaagtgaagatgagagataatcctaaactaaaaggaaaacct
gtcatcggttggcggtaaagcgagtcacgagcgtagtttctacggcatcatcacgaagca
20 agagcttatggtgttcactctgctatgcctatgactcaagcacataagctatgccccaat
ggatattatgtaacaagccgttttgatacttatagagaggtatctggtcaaactcatgaaa
atattcagaagttatacagaattagtagaaccatgtctttagatgaagcttatttagat
attacacatttagtgagaccggatttaccagcatcaaccattgcaaattatattcgaga
gatataacgaagtaaacacgtttaactgctcagctggcgtgtcttataataagttttta
25 gcaaagtttagcgagtggtatgaacaagccgaatggtttgacagtaattgattacaataat
gtacatgaaatattaatgcaattagatattggagattttccaggggtaaaaagcatatag

Sequence 3052

MDYFFAQVEMRDNPKLKGKPVIVGGKASHRGVVSTASYEARAYGVHSAMPMTQAHKLCPN
30 GYYVTSRFDYREVSGQIMKIFRSYTELVEPMSLDEAYLDITHLVRPDLASTIANIYIRR
DIYEVTRLTASAGVSYNKFLAKLASGMNKPNGLTVIDYNNVHEILMQLDIGDFPGVKSI*

Sequence 3053

35 Contig_0593_pos_608_1306

is similar to (with p-value 1.0e-57)

>sp:sp|Q41364|SOT1_SPIOL 2-OXOGLUTARATE/MALATE TRANSLOCATOR
PRECURSOR. >gp:gp|A47930|A47930_1 Sequence 1 from Patent WO
9534654. NID: g2301793. >gp:gp|U13238|SOU13238_1 Spinacia ol
40 eracea envelope membrane 2-oxoglutarate/malate translocator
(SODi1) mRNA, chloroplast mRNA encoding chloroplast protein
, complete cds. NID: g595680.

atggccttttttcatttcaagaggatttgtaaaaacagggctaggtcgacgtattgctctg
caattcggttaaattatttggaaagaaaacgcttggttggcttattcacttggtggtt
45 gaacttatcttagctcctgctacgccaagtaatacagcacgtgctggtggtattatggtt
ccaatcattaagtccttgctcagagtcatttgggttcacgcccagagatggttctgagaga
aaaatgggtgctgtttttaaacttttactgagttccaaggttaatttacttctcagctatg
tttttaacagctatggccggttaaccctatagcgcaaagtttagctgaaaaaacggcacac
gttcaaattacatggatgaattgggttgggtgctgctattataaccggattgatttctctc
50 atcgttgtccctttcattattttataaattataccacactactgttaaagaaacgcctaac
gctaaaaattggggtactgaacaactagaagaatgggacatatgtctatagccgaaaaa
ttgtatggttggtgcttttatcatagcattggcttgggttattaggaagcttcattaat
gttgatgccacgctcactgcatttattgcttttagcattgttactattaacaggtgtatta
gcgtggtcagatatttttaaatgaaacagatttctgcctaa

55

Sequence 3054

MAFFISRGFVKTLGRRIALQFVKLFGKKTGLAYSLVGVLDLILAPATPSNTARAGGIMF
PIIKSLSESGSPRDSERKMGAFLIFTEFQGNLITSAMFLTAMAGNPQAQSLAEKTAH
VQITWMNWFVAIIIPGLISLIVVPFIIYKLYPPTVKETPNAKKWATEQLEEMGHMSIAEK

LMVGVFIIALALWVLGSFINVDATLTAFIALALLLTGVLAWSDILNETDSA*

Sequence 3055

Contig_0593_pos_6260_5115

5 is similar to (with p-value 4.0e-36)
>sp:sp|P37469|DNAC_BACSU REPLICATIVE DNA HELICASE (EC 3.6.1
.-). >gp:gp|D26185|BAC180K_4 B. subtilis DNA, 180 kilobase r
egion of replication origin. NID: g467326. >gp:gp|Z99124|BSU
B0021_149 Bacillus subtilis complete genome (section 21 of 2
1): from 3999281 to 4214814. NID: g2636442.
10 atgttctcacatgatggtatgaagtcatttatggaatatgtattcgaagtcggttaaggta
gatcataacgaaatctatttaaaaaccacaaaagataagtcattcctagatatggacacc
atttcaaatttgtataactcaaaatttataggttacggattctttgaaagatatcaacaa
gatttgctcaatctttatcaaatagagcgtagcgaacgtattacaagaattcaattct
15 gatccgaatatacaaaattttgatgaaatgcttaacaaactacaaaaggtcagtttaatt
agtgaagtgaagaagtgggactaaaaaaattgtagatcactttgtcgaagaattatat
agcgaagaaccacaaaacaaaaatcaatacaggttataaaactgggtgattacaaaataggt
ggtttagaacctacacagttgattgtaatcgctgcgagaccgtcagtaggtaaaacgggg
tttgcgcttaatatgatgcttaatatagcgtctcaaggctataaaacttcattcttcagt
20 ctagagacaactggcggtgtctgtattgaaaaggatgttatcagcagaaactgggtagaa
ctaactcgtatcaaagaaattaaagatttagaacggatgaattaacacgtttaacaact
gcagcagacagaataactcaaacttgatatagatatcacgataaaaagcaatattactaca
catgatgtacgtaaacagcgatgaagaacaaagatgtgcaacaggttatcttcattgac
tacttacaacttatgcagacagacagtaagttagatcgctcgtaattggtatcgaaaagata
25 tcgagagatttgaagattattgcaaatgaaacaggtgcaattattgtgttgctatctcaa
ttgagcagaggtgtagaaacaagaaatgacaaaagacctatgctatctgacatgaaagaa
gcaggtggaattgaagcagatgcaagtttagctatgttggatatcgagatgattactac
aaccgtgatgatgttgatgactcaggcaagtcattgttgaaatgtaacatcgcaaagaat
aaagacggagaaacaggtgtagttgagtttgagtactacaagaaaacgcagaggttcttc
30 acatga

Sequence 3056

MFSDHGMKSFMEYVFEVGVKVDHNEIYLKTTKDKSFLDMDTISNLVNSKFIGYGGFFERYQQ
DLLNLYQIERTQNVLQEFNSDPNIQNFDEMLNKLQKVSLSASEESGTTKIVDHFVEELY
35 SEEPKQKINTGYKLVDYKIGGLEPTQLIVIAARPSVGKTGFALNMLNIASQGYKTSFFS
LETTGVSVLKRMLSAETGIELTRIKEIKDLEPDELTRLTAADRILKLDIDIDHDKSNITT
HDVRKQAMKNKDQVIFIDYQLMQTDSKLDRRNGIEKISRDLKIANETGAIIVLLSQ
LSRGVETRNDKRPMLSDMKEAGGIEADASLAMLRYDDYNNRDDVDDSGKSIVECNIAKN
KDGETGVVEFEYKKTQRFFT*

40

Sequence 3057

Contig_0594_pos_2706_2050

is similar to (with p-value 2.0e-76)
>sp:sp|Q59935|MANA_STRMU MANNOSE-6-PHOSPHATE ISOMERASE (EC
45 5.3.1.8) (PHOSPHOMANNOSE ISOMERASE) (PMI) (PHOSPHOHEXOMUTASE
) . >gp:gp|D16594|STRPMI_2 S.mutans pmi gene for mannosephosp
hate isomerase (complete cds) and scrK gene for fructokinase
(partial cds). NID: g451214.
atgagtatttttagttatttgagcaaatggaggcgtaggttctaaactagtaagtcaatta
aatgaagaacacggttgattttacagctgggtgtacgtaaggaagatcaagttaaagaatta
50 gaaaataaagggtattaaagctatattaatagatgtagaaaaaaatagtattaatgattta
aaaaatatctttacagattatgataaagttatcttttcagttggatctggtggaagcact
ggagcggataaaaacaatcattgttgatttagatgggtgctgtaaaaaacaattaaagctagt
aaagaagcgggtatcaaacattatgttatggtatcaacatacgattctagacgtgaagca
55 ttcgatgcgagtgagatttaaaacggtatatacaattgcaaagcattatgctgatgattac
ttaagaacatcagatcttaattatacaattgtacatccaggttcacttacagatgatgct
ggaactggaaaaatagaagctgatttatatttcgacaaagcaggatcaattccacgtgaa
gatgttgctacagtttttaaaagaagtagtaacttctgatggttttaataatcaagaattc
caaattttaagtggcaatcatggtgttaaagatgcattgaaaaactatgaatcttaa

Sequence 3058

MSILVIGANGGVGSKLVSQLNEEHVDFTAGVRKEDQVKELNKGKIKAILIDVEKNSINDL
KNIFTDYDKVIFSVGSGGSTGADKTIIVDLDGAVKTIKASKEAGIKHYVMVSTYDSRREA
5 FDASGDLKPYTIAKHYADDYLRSDLNNTIVHPGSLTDDAGTGKIEADLYFDKAGSI PRE
DVATVLKEVVTSDFGNNQEFQILSGNHGVKDALKKNYES*

Sequence 3059

Contig_0594_pos_1408_470

10 is similar to (with p-value 3.0e-28)

>sp:sp|Q04304|YMY0_YEAST HYPOTHETICAL 24.9 KD PROTEIN IN RC
A1-NPL6 INTERGENIC REGION. >pir:pir|S54466|S54466 hypothetic
al protein YM9582.15 - yeast (Saccharomyces cerevisiae) >gp:
gp|Z49259|SC9582X_15 S.cerevisiae chromosome XIII cosmid 958
15 2. NID: g807956.

atgccgttatttttaaaacccatttttctggataaagtatggggcagtgataatcttcgt
caatttgggtatcaactacctaataatcacatagtggaatggtgggaatttcagcacat
ccacacggaaaaagtggtgattgaaaatgggtatatttgcgtggcacaacattggatcaagta
tggacaatcatagagaaaatatttggagattttccaagtaaagattttccattaatggct
20 aaaattgtagatgctgctgcgccattgtctattcatgtacatcccgacgattcttatgca
tatgaacacgaagaaggtcaatatggaaaatctgaatggttggtacatcattgaagctgat
gaaggtgcaaagattactataggtacgtatgcgaaatctcgtgatgaatttgaagagcaa
ttggagcaaggtacatttgaaaattatttgagaacaatacaagtgcaaccaggtgatttt
tactttataccagctggaacgatacattccataggtgcaggcattatggcgtatgaagtc
25 atgcaatcatcagatatttcatatagaatttatgattatcatagaaaaactgataatagt
gaggaacgtgaattaaatatagataaggcattagatgttattaattattcaaatgaacta
cctaatatcactcctcaaaatgaagtgatagaaaatcacaattgtacacatattgtatct
agtgtatttttactatggttaagtgggatatttctggtactctaaattatatgaagcct
agagaattttgtcttctgtttctgttttagatggacaaggttaaacttattgtagatggtgat
30 atatatgatatatctaaaggttcaaaccttctgtgttaacttctgaagatttagatagtggt
ttcgaaggagatttttaactaatcattagttacatttaa

Sequence 3060

MPLFLKPIFLDKVWGSNDLNRQFGYQLPNNHIGECWGISAHPHGKSVIENGIFAGQTLQDV
35 WNNHREIFGDFPSKDFPLMAKIVDAAAPLSIHVHPDDSYAYEHEEGQYKSECWYIIEAD
EGAKITIGTYAKSRDEFEEQLEQTFENYLRTIQVPGDFYFIPAGTIHSIGAGIMAYEV
MQSSDISYRIYDYHRKTDNSEERELNIDKALDVINYSNELPNITPQNEVIENHNCTHIVS
SDFFTMVKWDISGTLNMYMKPREFCLVSVLDGQGKLIVDGDIYDISKGSNFVLTSEDLDV
FEGDFKLIISYI*

40

Sequence 3061

Contig_0595_pos_2494_3402

is similar to (with p-value 2.0e-60)

>sp:sp|P44331|RBSK_HAEIN RIBOKINASE (EC 2.7.1.15). >pir:pir
45 |B64073|B64073 ribokinase (rbsK) homolog - Haemophilus influ
enzae (strain Rd KW20) >gp:gp|U32732|U32732_6 Haemophilus in
fluenzae Rd section 47 of 163 of the complete genome. NID: g
1573480.

gtgattgtaattggatcaacaaatgtagataaaatttcttaattgtaaaagggtttccaaaa
50 cccggtgagacattacatatttaataagctcaaaaggagtttggtgggggcaaggagcc
aatcaagccatagcagctagtagattagcagcagatacaacatttatcagtaaagttggt
aaagatggcaatgccaaactttatattggaagatttcaaaaaagcaggtattcatacacia
tatatttttaacttcagaaaagtgaagaaactgggcaagcatttatcactgttgatgaagca
ggacaaaatacagattcttctgttacggtggtgcgaatatgacattaagtgcgaactgatgtt
55 gagatgagtgcggtatgcgtttattggtgcagactttgtgttagcgcagcttgaagttcca
tttgaggcgatagaacaagcatttaaaattgctcgttaaacaataatcactactgtatta
aatcctgcaccggcaattgaattgcctaagtcacttttagagtttaactgatataattatt
ccaaacgaacggaagcagaatttataacaggtatttcaatcaataatgaaagtgatag
aaagaaacagacatatattttctcgatttaggtatatctgcagatttaattacttttaggg

gagcaaggcacgtattgtgcatatcaagaacaatacaaaatgattcctgcgtgtaatgta
 aaagcaatagatacgacagcagcaggagatacatttataggtgcttttttaagtgaagta
 aataaagatttgagtaatatagaatcggtattcgacttgcaaatcaagcgtcgtctcta
 acggtacaacgaaaaggagcacaagcttctataccaacacgtaaagaagtagaggcgagaa
 5 tataattaa

Sequence 3062
 VIVIGSTNVDFKFLNVKRFKPGETLHINQAQKEFGGKGANQAIASRLAADTTFISKVG
 KDGNAFILEDFKKAGIHTQYILTSESEETGQAFITVDEAGQNTILVYGGANMTLSATDV
 10 EMSADAFIGADFVVAQLEVPFEAIEQAFKIARKQNITTVLNPAPAIELPKSLELTDIII
 PNETEAEELLTGISINNESDMKETATYFLDLGISAVLITLGEQGTTCAYQEYKMIACNV
 KAIDTTAAGDTFIGAFLSELNKDLSNIESAIRLANQASSLTVQRKGAQASIPTRKEVEAE
 YN*

Sequence 3063
 Contig_0596_pos_4550_3993
 is similar to (with p-value 3.0e-45)
 >sp:sp|P20368|ADH1_ZYMMO ALCOHOL DEHYDROGENASE I (EC 1.1.1.
 1) (ADH I). >pir:pir|A35260|A35260 alcohol dehydrogenase (EC
 20 1.1.1.1) I - Zymomonas mobilis >gp:gp|M32100|ZMOADHA_1 Z.mo
 bilis alcohol dehydrogenase I (adhA) gene, complete cds. NID
 : g155570.
 gtgagtggtatcgaaccaggacaatgggttaggcgtatttggtgtgggaggattaggtaat
 ttagcattgcaatacgccaaaaacgtaattggcgcgcaagtcgttgacattgacattaat
 25 gatgataaattaaattttgctaaagagcttggtgctgatgcaatcataaattcaactaat
 gttgatcctattgaggaagttaatcgtcctaacgaataataaaggattagatgcaacggtg
 attactgctgtagctaaaaacaccttttaatacaagcagttgatgttgttaaggcgggtgca
 cgtgtggttagcagtaggacttccagtagataaaatggatttagatattccacgttttagtg
 cttgatggaattgaagtcgttggttcattagttggtaccagacaagatttaagagaagca
 30 tttcaatttgctgccgaaaataaagttattcctaaaatccaattaagacaattatctgaa
 attaacgatatttttgatgaaatggaaaaaggaacaattacgggtcgaatggtaattgat
 atgaaaagcacgcactga

Sequence 3064
 VSGIEPGQWLGVFGVGGGLGNLALQYAKNVMGAKVVAFDINDDKLNFAGELGADAIINSTN
 35 VDPIEEVNRLTNNKGLDATVITAVAKTPFNQAVDVVKAGARVVAVGLPVDKMDLDIPRLV
 LDGIEVVGSLVGTQRDLREAFQFAAENKVIPIQLRQLSEINDIFDEMEKGTITGRMVID
 MKSTH*

Sequence 3065
 Contig_0598_pos_4948_5679
 is similar to (with p-value 3.0e-65)
 >sp:sp|P18843|NADE_ECOLI NH(3)-DEPENDENT NAD(+) SYNTHETASE
 (EC 6.3.5.1) (NITROGEN-REGULATORY PROTEIN). >gp:gp|D90817|D9
 45 0817_9 E.coli genomic DNA, Kohara clone #326(39.1-39.4 min.)
 . NID: g1742837. >gp:gp|D90818|D90818_4 E.coli genomic DNA,
 Kohara clone #327(39.2-39.5 min.). NID: g1742849. >gp:gp|AE0
 00269|AE000269_3 Escherichia coli K-12 MG1655 section 159 of
 400 of the complete genome. NID: g1788033.
 atgctaatagataaagcaagatcatttattcagaccatgtatagcgaattaaaatataat
 actaatgaaattgaaatagaatgaaagagattgagcaagaaattaaacttgactggtagt
 50 tacacacatacttattgaagaattatcttacggtgcacaaatggcatggagaaactcaa
 cgttgtattggttagactgttttgaattcttttaaatgttaaagatgcccagatgtatgt
 gacgaaaaagaattttataaaatttatacatcacatattaaagaagctactaacggcgga
 55 aaatcaaaccatatattacaatttttagtcctgaagatacacctaaaatttataataat
 cagttgatctggttatgctggttatgaaaatgttggcgatccatctgaaaaaagggttact
 cgtttagctgaacatctaggttggaaggtaaaaggttcaaattttgatattttacctctg
 atttatcaattgcctaacgacactataaaaatacacgaacttccaaatgatattgtaaaa
 gaagtttctatacatcatgaacactatcccaagctttcaaaattagggtttaaaatggtat

gcggtacctattattttcaaatatggatttaaaaatcggtgggtattacttaccctacagca
ccttttaattggatgggtatatggtaaccgaaattgctgtacgtaatttcacagacacctat
cgtcttatgtaa

5 Sequence 3066
MLIDKARFQITMYSELKYNTNEIENRMKEIEQEINLTGSYTHTYEELSYGAKMAWRNSN
RCIGRLFWSNLNVKDARDVCDEKEFIKFIHTHIKEATNGGKIKPYITIFSPEDTPKIYNN
QLIRYAGYENVGDPSEKKVTRLAEHLGWKGKGSNFDILPLIYQLPNDTIKIHLPNDIVK
EVSIIHHEHYPKLSKLGLKWYAVPIISNMDLKIGGITYPTAPFNGWYVTEIAVRNFTDTY
10 RLM*

Sequence 3067
Contig_0598_pos_4756_3287
unknown

15 gtgtaccaatataacgacgatagcttaatgttacacaatgatttatatacaaattaatatg
gctgaaagctactggaatgatggatccatgaaagaatagcagtggttgatttgattttt
cgaaaaatgccatttaataagtgatatgacggtattcaacggattgaaacgcgttgatgaat
ttcatcgaaaaactttgggtttacaaaatgaagatatcacatatttaaaatcgatagggtat
gaagaagattttctaaattacctaagaatttgaaatttacaggggaatattaaatctatg
20 caagaagggtgaaatttggttttgtaatgagccattattaagagttgaagcacctttaatc
caagcgcaacttattgaaactattttgttaaatatcattaatttccaaacattaattgca
actaaagccagccgaattcgtcaaatagcaactcatgacactttgatggaatttggtaca
agaagagctcaagagatcgatgctgcactgtggggcgctagagcagcctttattggaggg
tttgattctacaagtaattgttagagcaggaaaaacttttaataatacctgtatctggcaca
25 catgcacacgcactagtacaaacatatggtgatgagtatatagcattcaaaaagtatgct
gagcgacataaaaattgtgtgttcttagttgatacttttcatactttaaaatcaggagta
ccaaccgcaattaagggttgcaaaagagttaggagatactattaattttataggtatcaga
ttagattctggtgatattgctgacatctctaaagaagctcgtagaatgttagatgaggct
ggttttacagaagctaaaattatcgcatcaaatgatttggatgagcagactattacaagt
30 ttaaaagcacaaggcgctaaagttgacggatggggagtaggtacaaaactgattacagga
tatgatcaaccagccttaggtgcagtttataaaattgggttctctattgaaacagatgatggc
acaatgagtgatcgcatataaattatcaaaataatgctgagaaagttactacaccaggcaaa
aaaaatgtttatcgattattataataataaaacaggcaaggctgagggcgactatattacg
ctagaagggtgaaaatcctaataagcgaatctccattgaaaatgttccatcctgttcacact
35 taaaaatgaagtttattaaatcatttaaagcggttaattctacatcaatctatattgaa
aatggcaaaacttgataccatcttcagatgaatatgaagctcaggactatcttaaaaat
aatttaagtattttatgggaagaaaataaacgatattcttaaccgcaagattatccagta
gatttaagcactaaatgttgggaaaataagcataagcggtatttttgaagttgctgaacac
gttaaagagatggaggatgaaaatgagtag

40 Sequence 3068
VYQYNDDSLMLHNDLYQINMAESYWNDDGIHERIAVFDLYFRKMPFNNSGYAVFNGLKRVVN
FIENFGFTNEDITYLKSIGYEEDFLNLYKDLKFTGNIKSMQEGEICFGNEPLLREAPLI
QAQLIETILLNIINFQTLIATKASRIQIATHDTLMEFGTRRAQEIDAALWGARAFIGG
45 FDSTSNVRAGKLFNIPVSGTHAHALVQTYGDEYIAFKKYAERHKNCVFLVDTFHTLKSGV
PTAIKVAKELGDTINFIGIRLDSGDIAYLKSKARRMLDEAGFTEAKIIASNDLDEQTITS
LKAQGAQVDGWSGVGKLTIGYDQPALGAVYKLVSIETDDGTMSDRIKLSNNAEKVTPGK
KNVYRIINNKTGKAEGDYITILEGENPNDESPLKMFHPVHTYKMKFIKSFKAVNLHQSI
NGKLVYHLPDEYEAQDYLNKLSILWEENKRYLNPQDYPVDLSTKCWENKHKRIFEVAEH
50 VKEMEDENE*

Sequence 3069
Contig_0598_pos_3057_2467
is similar to (with p-value 5.0e-55)

55 >pir:pir|A47501|A47501 nitric-oxide synthase (EC 1.14.13.39
, endothelial - human >gp:gp|M93718|HUMNIOXSYN_1 Human nitric
oxide synthase mRNA, complete cds. NID: g189211. >gp:gp|L
10709|HUMNITOX17_1 Human constitutive endothelial nitric oxide
synthase gene, exons 25 and 26 and complete cds. NID: g34

8235. >gp:gp|M95296|HUMNOS_1 Human nitric oxide synthase mRNA
A, complete cds. NID: g189259. >gp:gp|L26914|HUMNOSA_1 Human
nitric oxide synthase mRNA, complete cds. NID: g434699.
gtgaaattaccttatgggtgtgcaacaagacgctcatgaagtagaagatgcacttgagttt
5 attaatcctgacacaacatatacagtttaattataaaccagcagttgatcagagtgttcaa
tcaacttagtgaaagcaggcatcaaacttactgattttcaaaagggtaatgaaaaagcacgt
gaacgaatgaaagtccaattttcaattgcttctaatactcaaggtatagtttttaggaact
gatcactctgccgaaaatattacaggattttacactaaatatggagatgggtgctgcggac
attgcgccctatctttgggttaataaaaagacaaggtaaacaattactagcttatctagga
10 gcacctaacaacacctttatgaaaaagtgccaacagctgatttagaagatgataaacctcag
ttaccagacgaggaagcactaggcgtatcttatcatgatattgatgattatttagaaggt
aaagaaattcctgcaactgctcgtgaaacaatcgaaaaacattatgtagaaatgcacat
aagcgtgaacttgcttatacacgatattcatggcctaataataacaaatga

15 Sequence 3070
VKLPYGVQQAHEVEDALEFINPDITYTVNIKPAVDQSVQSLSEAGIKLTDFQKGNEKAR
ERMKVQFSIASNTQGIIVLGTDHSAENITGFYTKYGDGAADIAPIFGLNKRQGGKQLLAYLG
APKHLYEKVPTADLEDDKQPLPDEEALGVSYHDIDDDYLEGKEIPATARETIEKHYVRNAH
KRELAYTRYSWPKYNK*

20 Sequence 3071
Contig_0605_pos_6367_0
is similar to (with p-value 2.0e-46)
>sp:sp|Q43157|RPE_SPIOL RIBULOSE-PHOSPHATE 3-EPIMERASE PREC
25 URSOR (EC 5.1.3.1) (PENTOSE-5- PHOSPHATE 3-EPIMERASE) (PPE)
(RPE). >gp:gp|AF070941|AF070941_1 Spinacia oleracea ribulose
-phosphate 3-epimerase (RPE) mRNA, nuclear gene encoding chl
oroplast protein, complete cds. NID: g3264787. >gp:gp|L42328
|SPIR5P3E_1 Spinacia oleracea nuclear-encoded chloroplast ri
30 bulose-5-phosphate 3-epimerase mRNA, complete cds. NID: g116
2979.
atgggttaaaattttaccatcactttttatctatagatttttttaaatttaaaagaagagctt
caattgttagaaacagcaaaggtagacggattacacttttgatgtaatggacggtaaattt
gtccctaattatttcaatcggatttccgatttttgatgctgttagacaacaatctcatttg
35 ccaatagatgttcatttaattgattgagcaacctgaaaattatattaatctttttgccgaa
catgggtgctgatatgatttctgttcatgttgagtcgacaacacatatacatagagcaatt
gaacaaattaaacaattagggaacaaagcaggtgtcgtcatcaatcctggaacatctgta
gaaacaattttacctatattgagtattgttgattatgttctagtaattgactgtaaatcct
ggttttggtggacaaacattcatagaacaatgcgtgactaagattgagcaattaaatcaa
40 cttaaacatgaaaatcatttaacttttgatattgaggtagatggaggcattaacgatcaa
acgagtaaacgatgtgtaga

Sequence 3072
45 MVKILPSLLSIDFLNLKEELQLLETAKVVDGLHFDVMDGKFVPNISIGIPIILDAVRQQSHL
PIDVHLMIEQPENYINLFAEHGADMISVHVESTTHIHRAIEQIKQLGKKAGVVINPGTSV
ETILPILSIVDYVLVMTVNPFGGGQTFIEQCVTKEQLNQLKHENHLTFDIEVDGGINDQ
TSKRCVX

Sequence 3073
50 Contig_0607_pos_1070_726
is similar to (with p-value 1.0e-60)
>pir:pir|I67760|I67760 transposase (insertion sequence IS10
) - Escherichia coli >gp:gp|S67119|S67119_2 BST=somatotropin
...BST/beta-Gal fusion protein [Escherichia coli, LBB84, pla
55 smid pXT107, IS10L/R-1, PlasmidTransposonInsertionMutant, 3
genes, 1679 nt]. NID: g455674.
atgcagattgaagaaaccttccgagacttgaaaagtctgcctacggactaggcctacgc
catagccgaacgagcagctcagagcgttttgatcatcatgctgctaatacgccctgatgctt
caactaacatggttggttgcgggcgttcatgctcagaaacaaggttgggacaagcacttc

caggctaacacagtcagaaatcgaaacgtactctcaacagttcgcttaggcattggaagtt
 ttgcggtcattctggctacacaataacaagggaagacttactcgtggctgcaaccctacta
 gctcaaaattttattcacacatggttacgcttttggggaaattatga

5 Sequence 3074

MQIEETFRDLKSPAYGLGLRHSRTSSSERFDIMLLIALMLQLTCWLAGVHAQKQGWDKHF
 QANTVRNRNVLSTVRLGMEVLRHSGYTITREDLLVAATLLAQNLFTHGYALGKL*

Sequence 3075

10 Contig_0608_pos_9823_8729

is similar to (with p-value 3.0e-52)

>sp:sp|Q43157|RPE_SPIOL RIBULOSE-PHOSPHATE 3-EPIMERASE PREC
 URSOR (EC 5.1.3.1) (PENTOSE-5- PHOSPHATE 3-EPIMERASE) (PPE)
 (RPE). >gp:gp|AF070941|AF070941_1 Spinacia oleracea ribulose
 15 -phosphate 3-epimerase (RPE) mRNA, nuclear gene encoding chl
 oroplast protein, complete cds. NID: g3264787. >gp:gp|L42328
 |SPIR5P3E_1 Spinacia oleracea nuclear-encoded chloroplast ri
 bulose-5-phosphate 3-epimerase mRNA, complete cds. NID: g116
 2979.

20 atgataactgcagaaaaaaagaagaaaaacaaattcttacctaatttcgaaaaacaatcg
 atctactccttaagatatgacgagatgcaacaatggcttattgatcacggacaacaaaaa
 ttcagagcaaaacaaatttttgaatgggtataccaaaagcgtgtgaatactattgatgaa
 atgactaacctgtctaaagagttacgtcaaattctcaaagatcattttgcaatgacgaca
 ttgaccactgttgtttaaacagaagtaagatggaacaattaagtcttattttgaatta
 25 caagatgggtatactattgaaactgttttaatgagacatgaatatggaaattctgtctgt
 gtaacaacacaagtaggatgtagaattgggtgtacgttttgtgcttccactttggcgga
 ttaaagcgtaattagaggccggagagattgtctctcaagtattaactgtacaaaaggca
 ctagacgaaaacgaatgaacgtgtatcacaaattgtcattatgggcataggtgaacctttc
 gagaattatgatgaaatgatggattttcttaagaattgttaatgatgataacagtttaaat
 30 attgggtgcacgtcatattactgtatctacttcaggaattattccaagaatttatgatttt
 gccgaagaagatatacaataaaattttgcgtgtgagtcctcatgggtgctaaagacgaaata
 agatcaagattaatgcctatcaatcgtgcttataacggttgataagttatggaagctatt
 cgttattatcaagaaaagacaaatcgccgtgttacttttgaatatggattgtttggtggt
 gttaatgaccaacttgaacatgagagatttggcacatttaattaagaatctcaattgc
 35 caggttaatttaataaccagtttaaccatgtcccagaaagaaattatgtaaagacacaaaaa
 gatgataatttttaaatcagagaaggaattaaagagattaggaattaatgtcacaattaga
 cgtgagcaagggtcagatattgatgtcgtggtggacaattaagagcgaaggaacgacaa
 gtagaaaacgaggttaa

40 Sequence 3076

MITAEKKKKKNKFLPNFEKQSIYSLRYDEMQQWLIDHGQKFRKQIFEWLYQKRVNTIDE
 MTNLSKELRQILKDHFAMTTLTTVVVKQESKDGTIKFLFELQDGYTIETVLMRHEYGNVC
 VTTQVGCRICTFCASTLGLKRNLEAGEIVSQVLTQKALDETNERVSQIVIMGIGEPF
 ENYDEMMDFLRIVNDDNSLNIGARHITVSTSGIIPRIYDFAEEDIQINFAVSLHGAKDEI
 45 RSRLMPINRAYNVDKLMEAIRYYQEKTNRRTVFYGLFGGVNDQLEHARDLAHLIKNLNC
 HVNLIPVNHVPERNYVKPKDDIFKFEKELKRLGINATIRREQSDIDAACGQLRAKERQ
 VETR*

Sequence 3077

50 Contig_0608_pos_4794_4123

is similar to (with p-value 7.0e-63)

>sp:sp|P36979|YFGB_ECOLI HYPOTHETICAL 43.1 KD PROTEIN IN ND
 K-GCPE INTERGENIC REGION. >gp:gp|D90881|D90881_3 E.coli geno
 mic DNA, Kohara clone #428(56.8-57.0 min.). NID: g1799913. >
 55 gp:gp|D90882|D90882_2 E.coli genomic DNA, Kohara clone #429(
 56.9-57.2 min.). NID: g1799919. >gp:gp|U02965|ECU02965_1 Esc
 herichia coli K12 ORF384 gene, complete cds, and ORF337 gene
 , partial cds. NID: g493518. >gp:gp|AE000338|AE000338_3 Esch
 erichia coli K-12 MG1655 section 228 of 400 of the complete

genome. NID: g1788862.

atggttaaaattttaccatcactttttatctatagatttttttaaattttaaaagaagagctt
caattggttagaaacagcaaaggttagacggattacacttttgatgtaatggacggtaaattt
gtccctaataatttcaatcgggtattccgattttggatgctgtagacaacaatctcatttg
5 ccaatagatgttcatttaattgattgagcaacctgaaaattatattaatctttttgccgaa
catggtgctgatatgattttctgttcattggtgagtcgacaacacatatacatagagcaatt
gaacaaattaaacaatttagggaaaaaagcaggtgtcgtcatcaatcctggaacatctgta
gaaacaattttacctatatattgagattgtttgattatgttctagtaatgactgtaaatcct
ggttttggtggacaaacattcatagaacaatgcgtgactaagattgagcaattaaatcaa
10 cttaaacatgaaaaatcatttaacttttgatattgaggtagatggaggcattaacgatcaa
acgagtaaacgatgtgtagaacaggggtgctacaatgttagtcactgggttcatacttcttt
aaacaagaggattatgcaaaagtaactagaaactttatgtcatattacagttttatcata
tatcatttatag

15 Sequence 3078

MVKILPSLLSIDFLNLKEELQLLETAKVDGLHFDVMDGKFVFNISIGIPILDAVRQQSHL
PIDVHLMIEQPENYINLFAEHGADMISVHVESTTHIHRAIEQIKQLGKKAGVVINPGTSV
ETILPILSIVDYVLVMTVNPFGGGQTFIEQCVTKIEQLNQLKHENHLTFDIEVDGGINDQ
TSKRCVEQGATMLVTGSYFFKQEDYAKVTRNFMSSYYSFIYHL*

20

Sequence 3079

Contig_0612_pos_14901_13966

is similar to (with p-value 1.0e-57)

>gp:gp|AB009866|AB009866_17 Bacteriophage phi PVL proviral
25 DNA, complete sequence. NID: g3341907.

atgcacttatctaaaatattaaaacaaggtaaaatcaaagctggagaaccaatggctaaa
acaggtaattcagggtcaatggactactgggtccacacgtacacttccaagttgaaagaggt
cgccatgatgacatcacaaacagagggacagtaaacctgctaaatggctcaaaggtcac
gggtggtggaagttggtggtgtagtggttctgtaaacgcacgtagagcaattcaaagagca
30 caatctatttttaggtggacggttataaatcgtcttataattaccgaacaaatgatgagagt
gccaaacgtgagtgtaacttccaatcagatgcggttaataactgggacatcaacgcacaa
aaaggaacgccttctaaaggtatgttccaaatgattgaaccatcttttagagcatatgct
aaaccaggacacggaacatcttaaatccaactgacgaagctatatctgctatgcgttac
attgtaggttaagtgggttcctattatggggagttggagaagtgcatttaaactgctgga
35 gattatgcttatgctacagggcggttattaacactgctggattatataatttggcagaa
gatggataccctgagatagtaaatccctacagatccaagcagacaatcagatgcatgaaa
ttgttacatcttgctgcaagtaaaattagtggaataacagaaataaacgcacctaaccaa
ttacgtacacctaattgttactagtaatacagttgataatgcagaattactactacaaatg
atagaaaaatcaacagaaacaaataaacgtgttaatggaaatagcacgaagtaataaaact
40 attgaaaaacaaccgaaaggtttttcagaacgcgatgtaagtcaggcacaaggttcaagg
ttaagactcgctgcttatagccaggagggtttataa

Sequence 3080

MHLSKILKQKIKAGEPMAKTGNSGQWTTGPHVHFQVERGRHDDITNRGTVNPAKWLGKH
45 GGGKVGSGSVNARRAIQRAQSILGGRYKSSYTEQMMRVAKRESNFQSDAVNNWDINAQ
KGTPSKGMFQMIEPSFRAYAKPGHGNI LNPTDEA ISAMRYIVGKWVPIMGSWRSFAFKRAG
DYAYATGGVINTAGLYNLAEDGYPEIIVPTDPSRQSDAMKLLHLAASKISGNNRNKRPNQ
L RTPNVTSNTVDNAELLQMIENQQKQINVLMEIARSNKTIEKQPKGFSE RDVSQAQGSRL
50 LRLAAYSQGGL*

50

Sequence 3081

Contig_0612_pos_6018_3844

>pir:pir|S01788|S01788 formate C-acetyltransferase (EC 2.3.
1.54) - Escherichia coli >gp:gp|D90728|D90728_3 Escherichia
55 coli genomic DNA. (20.3 - 20.7 min). NID: g1651424. >gp:gp|X
08035|ECPFL_1 E. coli pfl gene for pyruvate formate-lyase (E
C 2.3.1.54). NID: g42369. >gp:gp|AE000192|AE000192_6 Escheri
chia coli K-12 MG1655 section 82 of 400 of the complete geno
me. NID: g1787125.

gtgagagaattcatacaattgaactattcattatatgaaggtgacgatgaatTTTTAGAA
 ggTcctacaaaagcaactgaaactttatgggatcaagtaatgcaattatcaaaagaagaa
 cgtgagcgcgggtggcatgtgggacatggatactaaagtggcatcgacaatcacttctcat
 gacgttgggttacttagacaaagatttagaaaaagttgttggtgttcaaaactgaaaaacca
 5 ttcaaacgttctatgcaaccattcgggtggtattcgtatggcaaaagcagcatgtgaagcg
 tatggttacgaatttagatccagaaacagaaaaaatcttcaactgaatatcgtaaaacacac
 aaccaaggtgtattcgatgcatattcaagagaaatgttaaactgtcgtaaagctgggtatt
 attactgggtttgccagatgcttacggacgtggacgtattatcggagactatcgctcgtgtt
 gctttatacgggtgtagatttcttaattggaacaaaaacttaaagactttaacacaatgtct
 10 actgaaatgtctgaagatgtaattcgtttacgtgaagaattatcagagcaatatcgttca
 cttcaagatttaaagaattaggtcaaaaatatggatttgatattagccgtcctgctact
 aacttcaaagaagctgtgcaatgggtatacttagcatatttagctgctatcaaagaacaa
 aatgggtgcagcaatgagtttaggacgtacttcaacattcttagatatttatgctgaacgt
 gatttacaaaatggtgacatcactgaacaagaagttcaagaaatcattgaccacttcatt
 15 atgaaattgcgtatcggttaaattcgcgcgtacgcctgaatataatgaattattctctgga
 gatccaacttgggtaactgaatctatcgggtggtgtaggtattgacggacgtccaatggta
 actaaaaactcattccgtttcttacactcattagataatttaggtccagcaccagaacca
 aacttaacagtggttatgggtcactcgttacctgaaaacttcaaaatctatttggtctaaa
 atgagtatataaacgagctcaatccaatatgaaaatgatgacttaatgcgtgaaagctat
 20 ggcatgattatgggtatcgcttgctgtgtatctgccatgaagattggtaaacaaatgcaa
 ttcttcgggtgcacgtgctaaacttagctaaagcattactttacgctatcaatgggtggtaaa
 gatgaaaaatctggtaaacaaagttgggccaagttatgaaggtattaaatcagacgtacta
 gattatgatgaagtcttcgaaagatatgaaaaatgatggactgggttagctggcgatatat
 atcaactcattaaatatcattcactatatgcatgataaatatagctatgaacgtcttgaa
 25 atggctttacatgatacagaaattattcgacaaatggcaactgggtattgccggattgtct
 gtagcagctgactctttatcagcgattaaatatgcacaagttaaacctatccgtaacgaa
 gaaggtcttgtaactgactttaaaatcgaaggcgacttccctaaatatggtaataatgac
 agtcgtgttgatgaaattgcagtagatttagttgaacgtttcatgactaaattacgtagc
 cataaaacataccgtaattctgaacacacaatgagtgatttaacaattacttcaaagctt
 30 gtttatggtaagaaaactggtaacacaccagatggacgtaaagctggcgaaccatttgca
 cctggcgcaaacccaatgcatgggtcgtgacccaaaaggtgcattatcttcaactaagttca
 gtagctaaaataccttacgattgctgtaaagatgggtatctcaaatacatttagtatcgta
 ccgaaatcactaggtaagaagaagcagatcaaaaataaaaacttaactagtatggttagat
 gggtatgcaatgcaacatgggtcatcacctcaacattaacgtatttaatagagaaacatta
 35 attgatgcaatggaacatccagaagagtatccacaattaacgattcgtgtatctgggatac
 gctgtaaacttcattaaattaacacgtgaacaacaattagatgttatttcacgtacattc
 cacgaatctatgtaa

Sequence 3082

40 VREFIQLNYSLEYGDDDEFLEGPTKATETLWDQVMQLSKEERERGGMWDMDTKVASTITSH
 DAGYLDKDKLEKVVGVQTEKPFKRSMQPFGGIRMAKACEAYGYELDPTEKIFTEYRKTH
 NQGVFDAYSREMLNCRKAGIITGLPDAYGRGRIIGDYRRVALYGVDFLMEQKLKDFNTMS
 TEMSEDEVIRLREELSEQYRSLQDLKELGQKYGFDISRPATNFKEAVQWLYLAYLAAIKEQ
 45 NGAAMSLGRTSTFLDIYAERDLQNGDITEQEVQEIIDHFIMKLRIKVFARTPEYNELFSG
 DPTWVTESIGGVGIDGRPMVTKNSFRFLHSLDNLGPAPEPNLTVLWSTRLPENFKIYCAK
 MSIKTSSIQYENDDLMRESYGDDYGIACCVSAMKIGKQMQFFGARANLAKALLYAINGGK
 DEKSGKQVGPSEYEGIKSDVLDYDEVFEREYKMDWLAGVYINSLNIIHYMHDKYSYERLE
 MALHDTEIIRTMATGIAGLSVAADSLSAIKYAQVKPIRNEEGLVTDFKIEGDFPKYGNND
 50 SRVDEIAVDLVERFMTKLRSHTYRNSEHTMSVLITSNVYVGKKTGNTPDGRKAGEPFA
 PGANPMHGRDQKALSSLSSVAKIPYDCKDGISNTFSIVPKSLGKEADQNKNLTSMLD
 GYAMQHGHLNINVFNRETLIDAMEHPPEYPQLTIRVSGYAVNFIKLTREQQLDVISRTF
 HESM*

Sequence 3083

55 Contig_0613_pos_918_1343
 is similar to (with p-value 1.0e-32)
 >gp:gp|AB009866|AB009866_61 Bacteriophage phi PVL proviral
 DNA, complete sequence. NID: g3341907.
 atgacaaatacattagaaattaaattattatcagaaaacgcgactatgccgaagagagca

5 aattctacagatagtggttgacttatacgtatcagaaacgattaacattcctgcacac
gcaactaaagtagttaaaacagatatagcgattaatctgccttatgggtatgaggcgcaa
gtaagacctagatctggtaaatcacttaaaactaaattgcgtgtagcactaggaacaata
gaccaaacataaccacaaagaaataggtatcatcacagataacataggtaatgaagatc
acagtagaaaaaggagaaagattagcgcaattagttgtagcgccagttgtatatcctaca
cccaaacaggttgattggtttgaaaatgaaagcgacagaggtgcatatggaagcacagga
gaataa

Sequence 3084
10 MTNTLEIKLLSENATMPKRANSTDSGLDLVSETINIPAHATKVVKTDIAINLPYGYEAQ
VRPRSGKSLKTKLRVALGTIDQTYHKEIGIITDNIGNEDITVEKGERLAQLVVPVYPT
PKQVDWFENESDRGAYGSTGE*

Sequence 3085
15 Contig_0613_pos_2864_3259
is similar to (with p-value 2.0e-21)
>gp:gp|AB009866|AB009866_1 Bacteriophage phi PVL proviral D
NA, complete sequence. NID: g3341907.
gtgaaagaaatgatataactataagtggatgaccaatattataacatcacaaatttac
20 gatgcagacagtagatctattgcacaatatgggtatagaatctgttatgcctaaggctaaa
ggacaagcaggcgataaagtgtttttgaagggttatgaatagaaataaagcgtggagacgt
aatttaaaactcatagagaagattgagttttattgacaagtagtaggagtagacataactaac
gagatgaactttcatatattgcagatgattaagtttaggtacgaaacttaaaactgttatg
gacttaattggaaataaaaaagtaaatcaacattctatggttggtttaaatgaaatagtaaac
25 gtatataatggatgcacagaaaggatattacgattag

Sequence 3086
VKEMIYNYKWMNTIITSQIYDADSTSIAQYGIESVMPKAKGQAGDKVFLKVMNRNKAWRR
NLKLEIEKIEFIDKYEYITNEMNFHILQMIKLGTKLKTVMDLMEIKSKSTFYGCLNEIVN
30 VYMDAQKGYD*

Sequence 3087
Contig_0613_pos_4377_4844
35 >gp:gp|X97563|BPHA3GP3_5 Bacteriophage A2 gp3 gene and 4 op
en reading frames. NID: g1523807.
atgccgccaagaaaattattatctcaacaaaaaggttaatttaacagtcgaacaacaagaa
aataaagaaaatgcagaaaaagcgatggcgcaactcactgagatagatgaaaaacctcct
gaatggcttgataaagatgcaataaaagaatggcatcgcatattacctttgattcaagaa
ctaccaatagcagcttttagatatggggttattagccacctattgtcaaacatatagcaat
40 tacaagaacgccacgatccaattagaaaaagaggggtatgatcgctcgaaccgaaagagga
acgaaattatctagttattacacagtagcaaaagagatagcgtaaatgctatgaactccatt
tgtcctaaattaggattgacagttgagtcacgtttaaaaatattgtcgccagatacaaaag
aaagaaaaagaaagatgaatttgaggacttaataatggcaaagattag

45 Sequence 3088
MPPRKLSSQQKGNLTVEQQENKENAEKAMAQLTEIDEKPPEWLKDKAKEWHRILPLIQE
LPAAALDMGLLATYCYQTSNYKNATIQLKEGMIVETERGKLSYYTVQRDSVNAMNSI
CPKLGTLVESRLKILSPDTKKEKKDEFEDLMNGKD*

50 Sequence 3089
Contig_0613_pos_4909_6483
is similar to (with p-value 6.0e-23)
>sp:sp|P14597|DUT_ORF2 DEOXYURIDINE 5'-TRIPHOSPHATE NUCLEO
TIDOHYDROLASE (EC 3.6.1.23) (DUTPASE) (DUTP PYROPHOSPHATASE)
55 . >gp:gp|M30023|ORFPRTPS_1 Orf virus homologue of retroviral
pseudoprotease gene, complete cds. NID: g332561.
gtgaaagcctgtcaacgccatttagatgacttgaacgattcggaaactcccttatcatttt
gatgtaaaagaaagctaatacacattattaagtttcttgaaatgttgccagatcctaaaact
ggtaaacatttatcggttaggcggttttcaaaaattcatttgctggttagcttaaatggttgg

tacgacagacatgggtacaaaagatttacaaaagcctatatatcaatgagcagaaaaaat
 ggtaaaaacattattgatctctggaatggcattgtacgatttattgatgggtaaagatccg
 ttgaatgaacgggttgattgggtttgagcgccaattcaagagaccaagctgggtatagcatac
 gatatgacattggcacaactgaaaagctattagaagcgtttctcctaagggttaaatcgatg
 5 actaagataacgccaagtgcaaaagaaatattgaatattaatgatcgaagtaaagttaaa
 gccgtttcaaatgaagctgcaaatttagaaggctcatcagtttagctacgcaatcatcgat
 gaatatcatgaagctaaagataaaaagatttatgaaacggttaagacgtgggcaagtgc
 ctgcacaaccttatattaattattatctcaacagctggaaactaatttgatgggtccgatg
 tatgaagaatatttatataattgataagatacttgacggcatagcaaaaaatgaaaactac
 10 ttgttttctgtgctgaacaagatgatgagaaagaagtatatgacgttaaaaacttggtt
 aaatccaatccacttatggagttgccagaaatggcacaattgttaactaagaatattcaa
 ccagaagttaaaactgcaattgatagtggttcaggattaaatgggatattaataaagaat
 ttcaatatgtggcgtgcagcaagcacagaatcttatttagattttcaatgattggaagaaa
 aatgaaatagactttgatataaatggctctaaaacttatatcggttttagacttatcgctg
 15 gctgacgacttaaccgcagtatcggtttgttcatcttgatgaagataatcaagagtattat
 gtaactagtcattcggtttgtggtactaaagggtgatttagatggcaagattgatagagac
 ttatttgatttacagacaacttgacagaaagtggttattgtacgattaccgatttacaaagt
 ggaattatcaatactgaccaagttttaaatattacattgagaattatatcgaccaatataaaa
 ttagacgtacaagcgttatgttatgatccttactcaatacatgggtgtattgcagaaatt
 20 gagcgtatagattggccttatgatttagtagaaatcagacaagggccacaaacactatct
 aatccgatactggatttttagactgaaagtgattaatggggacatcaagcatcataaaaat
 ccgttactagacattgcagtcaaaaatgctgtggcaaaagataccaatgactcattaatg
 attgaaaagaagatgaaccgagaaaaaatagatccactcatgggtaccatatttgcttat
 gttataaatagctga

25

Sequence 3090

VKACQRHLDLNDSELPYHFDVKKANHIIKFLEMLPDPKTKGQLSLGGFQKFIAGSLNGW
 YDRHGYKRFTKAYISMSRKNKTLISGMALYDLMGKDPLNERLIGLSANSRDQAGIAY
 DMTLAQLKAIRSVSPKVKSMTKITPSAKEILNINDRSKVKAVSNEANLEGHQFSYAIID
 30 EYHEAKDKKIYETLRRGQVLLHNPILIIISTAGTNLNGPMYEEYLYIDKILDGIAKNENY
 FVFCAEQDDEKEVYDVKTWIKSNPLMELPEMAQLLTKNIQPEVKTAIDSGSGLNGILIKN
 FNMWRAASTESYLDNFNDWKKNEIDFDINGSKTYIGLDSLRADDLTAVSFVHLEDDNQEYY
 VTSHSFVATKGGLDGKIDRDFIDYRQLAESGYCTITDLQSGIINTDQVLNYIENYIDQYK
 LDVQALCYDPYSIHGVIAEIERRDWPYDLVEIRQGPQTLNPIILDFRLKVINIGDIKHHKN
 35 PLLDIAVKNAVAKDTNDSLMIIEKKMNREKIDPLMATIFAYVINS*

Sequence 3091

Contig_0619_pos_4899_6257

>sp:sp|069282|MQO_CORGL MALATE:QUINONE OXIDOREDUCTASE (EC 1
 40 .1.99.16) (MALATE DEHYDROGENASE (ACCEPTOR)) (MQO).
 gtgagtcacccggcggaagaaagttcaaatgtttggaataatgcgggaacagggtcattca
 gcattgtgtgaattgaactatacgaaagaaggtaaagatgggttcagtagatattactaaa
 gcaattcatatttaacgagcaatttcaaatatctaaacagttttgggcttatttaatacgt
 gaaggtcataattgaaagtcagataaaatttattcaatcagtgccacatatgagctttgtt
 45 aaaggggaagaaaaatgttaaatttttaaaaagtcgagtgaggagtttacagaaaaatgta
 ttatttgaaaaaatgaaaatttctcaagatccagaaaaaatcaactcatgggttccttta
 atgatggaaggacgccaatcagatgaagcaattgccattacgtatgacgagacagggtaca
 gatgttaactttgggtgctttgactaaaaagtttaataagctaatttacaacaaaaaatgtt
 ggcatttaattataaacatgaagtttttagatataaaaaaatataaatggttaactggcaa
 50 gttgtgggttaaagatttaaatatcatcaaatgtaattgaattatgaattcgaagtcgtcttc
 atcggagctgggtggtgcaagtttacctttattacaaaaaacaaagattaaggaatctaaa
 cacattgggtgggtttcccagtaagtggattatttttacgatgtaaaaatccagatgtcata
 catagacatcatgcaaaagtctacggtaaaagccgaggttggtgcacctccaatgtcagtt
 ccacatttagatacacgatttggttaattggtgaaaaatcattactatttggaaccttttgca
 55 ggggttttcgcaaaaattcttaaaaaacgggttcataatttagatttagttaaatctgtgaaa
 cccaataatattgataacaatgttaagtgtggcgtaaaagaatttaatttgacgaaatat
 ttagtttctcaattaatgcttttcaaatgaagaacgggatcaatgatttgcggtgattctta
 ccagaagcgaaagatgaagattgggaagtaattactgcaggtcaacgtgttcaagtaatt
 aaagatacacgataagtcataaagggtcaattacaatttggtacggaagtaataacatcagaa

gatgggttcacttgctgcattattaggtgcttcacctgggtgcttcgactgctgttgatatac
atgtttgatgtcttgcaacgttggttacaatcagagtttaagtcagggaaacaaaaaatt
aaagaaatgggtcccatcatttggtttaaaattgtcagagcatgaagatatgtaccattca
ataaacgaagaagtaaaaaatacttgaatgtaaagtag

5

Sequence 3092

VSQPGEESNVWNNAGTGHSA LCELNYTKEGKDGSDITKAIHINEQFQISKQFWAYLIR
EGHIESPDKFIQSVPHMSFVKGEENVKFLKSRVASLQKNVLFKMKISQDPEKINSWVPL
MMEGRQSDAIAITYDEGTVDVNFAGALTKKLIANLQQKNVGINYKHEVLDIKKLNNGNWQ
VVVKDLNTSNVMNYESKFVFIGAGGASLPLQKTKIKESKHIGGFVSGFLRCKNPDVI
10 HRRHAKVYGAIEVGAPPMSVPHLDTRFVNGEKSLLFGPFAGFSPKFLKNGSYLDLVKSVK
PNNMITMLSAGVKEFNLT KYLVSQLMLSNEERINDLRVFLPEAKDEDWEVITAGQRVQVI
KDTDKSKGQLQFGTEVITSEDGSLAALLGASPGASTAVDIMFDVLQRCYKSEFKSWEPKI
KEMVPSFGLKLSEHEDMYHSINEEVKKYLNVK*

15

Sequence 3093

Contig_0620_pos_2802_3113

is similar to (with p-value 7.0e-21)

>pir:pir|A00461|DEECR NADH dehydrogenase (EC 1.6.99.3) - Es
20 cherichia coli >gp:gp|D90746|D90746_3 Escherichia coli genom
ic DNA. (24.9 - 25.3 min). NID: g1651543. >gp:gp|V00306|ECNDH
X_1 E. coli gene ndh coding for respiratory NADH dehydrogena
se (a component of the electron transport chain). This enzym
e catalyses the transfer of electrons from NADH to the respi
25 ratory chain and thus links the major catabolic and energy-p
roducing pathways of the cell. NID: g42112. >gp:gp|AE000211|
AE000211_7 Escherichia coli K-12 MG1655 section 101 of 400 o
f the complete genome. NID: g1787345.

30

atgttaaaaaataacgacatgtctgatgggtattttaaaaattaaagtaaattgggtggagga
tgcacaggattaaacttatgggtatgtcagccgaagcagaacctgggtgaaaatgatgaaatt
tgcgaatactatgggtttgaaagttctagtagaccgaaatgatgctcctgtattaaatgga
acaacaattgatttttaaacagtcacttatgggtggaggatttcaaataaacaatcctaatt
gctattgcctcatgtggttgtggaagttcattttaaaacagctaaagtcgctggaaatcca
gagcaatgttaa

35

Sequence 3094

MLKNNDMSDGYLKI KIVNGGGCTGLTYGMSAEAEPEGNEDEILEYYGLKVLVDRNDAPVLNG
TTIDFKQSLMGGGFQINNPNAIASCGCGSSFKTAKVAGNPEQC*

40

Sequence 3095

Contig_0620_pos_3383_4507

is similar to (with p-value 9.0e-18)

>sp:sp|Q44540|YNIU_AZОВI HYPOTHETICAL 11.0 KD PROTEIN IN NI
FU 5'REGION (ORF6). >gp:gp|M20568|AVINIFC_16 A.vinelandii ma
45 jor nif gene cluster encoding nitrogen fixation complex, com
plete cds. NID: g758356.

50

atggctcaagatcgtaagaaagtgcctcgattaggtgcaggttatgctgggtttacaaact
gtaactaaattacaaaaagaactttctgctgatgcagcggaattactttaattaataag
aatgaatatcattatgaatcaacttggttgcatgaagcttctgccggtacgattaattat
gaagatttatgtattctgttgagaaaactgtcaacaaaaataaagtgaattttgtgtgt
gctgagggtaaacaaaattgatcgtaattgctaaacgtgtagaaaactgataaggggtgtttat
gactttgatatacttagttgttgactaggtttttagtagcgaacatttggtattgatggt
atgaaagaacatgctttccaaattgagaacgttttaacttctcgaagttgtctcgtcac
attgaagataagttcgcaatttatgctgcttctaaagaaaaagatgataaagatttatct
55 attttagttgggggagctggatttacaggaattgaatttctaggtgaattaactgataga
attcctgaattatgcagtaaatatgggtgttgatcaaagtaaagtgaagttaacatgtgtt
gaagcagcacctaaaatgttaccgatgttctcagacgacttagttagttatgcagtaaaa
tatttagaagaccgtggagtgaattcgaacacacattgtcgcttgtaaatgaa
aaaggtttcggtgttgtaagtcaatggagaaaaacaacaaattagaagccggaacttctgta

tggactgctggagtgcgtggaagtcatttaaatggaagaatcatttgaaggtgttaaactg
 ggacgtattatcaataaacaagatttaacaattgaaggtcataatgacatctttgttata
 ggagattgttcagcggtttattccagctgggtgaagagcgctccattaccaacaacagctcaa
 attgctatgcaacaaggtgagcactgctagcaacattaaacggtttattaaatgggtgaa
 5 tcaacacaagatttccaatatgttaacctaacgatcaagatgcagttgaaaggaaagaga
 aagaaactgagtttcaaaaacaacaagatgaagaaattgctttaa

Sequence 3096

MAQDRKKVLVLGAGYAGLQTVTKLQKELSADAAEITLINKNEYHYESTWLHEASAGTINY
 10 EDLLYPVEKTVNKNKVNFEVVAEVTKIDRNAKRVETDKGVYDFDILVVALGFVSETFGIDG
 MKEHAFQIENVLTSRKLSRHIEDKFANYAASKEKDDKDL SILVGGAGFTGIEFLGELTDR
 IPELCSKYGVDQSKVKLTCVEAAPKMLPMFSDDLVS YAVKYLEDRGVEFKIATPIVACNE
 KGFVVEVNGEKQQLLEAGTSVWTAGVRGSHLMEESFEGVKRGRIINKQDLTIEGHNDIFVI
 GDCSAFIPAGEERPLPTTAQIAMQQGEHTASNIKRLNLNGESTQDFQYVNLTIKMQLKGR
 15 KKLSPKNNKMKLL*

Sequence 3097

Contig_0620_pos_2198_1134

is similar to (with p-value 7.0e-29)

20 >gp:gp|Y09899|CVPME131_3 C.viguieri phS gene, gene encoding
 putative NADH dehydrogenase and two genes encoding unknown
 proteins. NID: g2765033.
 atgaaaaacttagtattactagggcggttatggtaatatgcgaattatgtcgcgcat
 ttacctcattcaattcctgagggatattcacttaactttaatcgaccgcatgccattccac
 25 ggtttaaaacctgaattttatgcacttgcagcaggaactaaatctgacaaagaggtgcga
 atccaatttccagatagcagtcataataacggtttatggggaaatcagtgatatagat
 ttggacgaacaaatgataacagttggaaattcaaaaatagattatgacgaacttatcatt
 ggtctaggggtgtgaggataaatatcataatgtccctgggtgctgaagcatatacacatagc
 attcaaacattatctaaatcgcgtaaacataccatagaattagcgagttacctaagggt
 30 gcacgtgtaggtatcggtggggcaggtttaagtggcattgaattagcgagcgagctacgt
 gaaagtcgatcagacttggaaattttgttatatgatagagggcctcgaattttaaggaat
 tttccagagaaactgagtaaatacatatctaatgggttttctaaacacaatgttactgta
 gtacctaatcagtcacgacagagtagaacccggaaaaattataataatggtaaacca
 gaaaatattgatttagtcgtttggacagcaggcatacaacctgttgaaattgtgcgtaat
 35 ctctctattgatatgagtaccactggacgcgtaattattaatcagtagcatcaagtcacca
 acctatagaaatgtttatgtcgtaggtgactgtgctaatttaccacatgcacccagtgct
 caactagcagaactacaaggtgaacagattgctgaggtgtgagaagcaatggaataac
 gaaccacttccagataaaaatgcctgaaattaaagtacaaggcttttttaggctcttttaggt
 gacaaacaaggttttgcttatatcatggatcgaacagttaccggacgattagcctctatt
 40 ctaaaatcaggtgttctgtggcgctataaatatcataatgggttaa

Sequence 3098

MKNLVLLGGGYGNMRIMSRILPHSIPEGYHLTLIDRMPFHGLKPEFYALAAAGTKSDKEVR
 IQFPDSSQINTVYGEISDIDLDEQMITVGN SKIDYDELIIGLGCEDKYHNVPGAEAYTHS
 45 IQTLKSRETYHRISELPKGARVGIVGAGLSGIELASELRESRSDLEILLYDRGPRI LRN
 FPEKLSKYISNWF SKHNVTVVPNSVIDRVEPGKIYNNGK PENIDL VVWTAGIQPVEIVRN
 LPIDMSTTGRVIINQYHQVPTVRNVVVGDCANLPHAPSAQLAELQGEQIAEVLKKQWNN
 EPLPDKMPEIKVQGFLGSLGDKQGFAYIMDRVTVTGRLASILKSGVLWRYKYHNG*

50 Sequence 3099

Contig_0622_pos_4948_4412

is similar to (with p-value 1.0e-61)

>gp:gp|M57622|BACRGB_1 B.stearothermophilus ribosomal prote
 in L6 gene, complete cds. NID: g143418.
 55 atgagtcgtgttggtgaagaaaattattgacattcctagtgacgtaacagtaacttttgac
 ggaagtcattgctactgtaaaagggtccaaaagggtgaattagaaagaacttttaaatgaaaga
 atgacatttaacaagaagaaacactgttgaagttgtaagaccatctgattctaaagaa
 gacagaacagatcatggtacaactcgtgctttattaaataatattggtactaggtgtttct
 caaggttacgaaaaaacacttgagcttgggtgtaggttaccgtgcacaaatgcaaggt

aaagatttagtacttaatggttgatactctcaccagttgaaattaaagcagaagaaggc
attactttcgctggttgagaaaaatacaactgttaaagtatctggtgtttctaaagaacaa
gttggtgcgattgcttcttaacattcggtctgtaagacctccagaaccttataaaggtaaa
ggtattcgctaccaaggtgaatatgtacgccgtaaagaaggtaaaactggtaataa

5

Sequence 3100

MSRVGKKIIDIPSDVTVTFDGSHTVKGPKGELERTLNERMTFKQEENTVEVVRPSDSKE
DRTDHGTTTRALLNNMVLGVSGQYEKLELVGVGYRAQMKGDLVLNVGYSHPVEIKAEEG
ITFAVEKNTTVKVSQVSGVSKQVGAIASNIRSVRPPEPYKKGKIRYQGEYVRRKEGKTGK*

10

Sequence 3101

Contig_0623_pos_10299_9358

is similar to (with p-value 7.0e-34)

>gp:gp|AF017231|AF017231_1 Trypanosoma brucei brucei inosin
e-adenosine-guanosine-nucleoside hydrolase mRNA, complete cd
s. NID: g2645494.

atgaccaaagtatatattttaatcatgatggtggcggttgatgatctagtgtcactattttta
ttattacaaatggaaaatatagaacttggtggcgtaagtacgattggtgcagactgctat
ttagagccttcattaagtgttcattaaagattataaatcggttttccagacgttgaaatt
aatgtagcaccatcatatgaaagagggaaaaatccttttccaaaagaatggagaatgcat
gctttctttatggatgccttaccagtgctaaatgagtccttgatatacccaaaagatgcaag
gctagtgaggacgagggcgatatagatattattcgtaaagtgaagagttgtgatgagaaa
gttacattgttatttactggaccgcttacagatttagctaaagctataaaatatgacaac
tcaatttttaaaaaatatagagaaattagtttggatgggtggaacgtttttagacaaagga
aatgttgaagaaccagaacatgatggtacagctgaatggaatgcattttgggacctgag
gctgtaaaagttgtattagatagtgatgatgaatgtcgatattggtgcttttagaaagtaca
aatcaagtcctcctaacaatggaagttcgtcaaatgtgggcagataaaagacaatattta
ggcggttgattttctgggcacaagttacgcagcagtagaccaccactcacacattttgtgacc
aattcaacatactttttatgggatgtattactactgcttatgtgggttctccaaattta
gttgaaatcaacgaaattgaaaattgatgtagtcagtcaggacctagtcaaggggagaaca
ttccaatctgaatatggacgtgaagttcaagtcattacggatgtaataaacaagcattt
tttaactacataacggatttagcaaagaaaatcgagtcctaa

Sequence 3102

MTKVYFNHDGGVDDLVSFLLLQMENIELVGVSTIGADCYLEPSLSASLKIINRFSDEVI
NVAPSYERGNPFPPKEWRMHAFMDALPVLNESCIPKRCASEDEAYIDIIRKVKSCDEK
VTLLFTGPLTDLAKAIKYDNSILKNIKLVWGGTFLDKGNVEEPEHDGTAEWNAFWDPE
AVKVVLDSDMNVDIVALESTNQVPLTMEVRQMWADKRQYLGVDLGTSYAAVPPLTHFVT
NSTYFLWDVLTAYVGSPLNVESTKLKIDVVSQGPSQGRFTFQSEYGREVQVITDVNKQAF
FNYITDLAKKIES*

Sequence 3103

Contig_0625_pos_635_1573

is similar to (with p-value 4.0e-70)

>sp:sp|P54524|YQIG_BACSU PROBABLE NADH-DEPENDENT FLAVIN OXI
DOREDUCTASE YQIG (EC 1.-.-.-). >gp:gp|D84432|BACJH642_230 Ba
cillus subtilis DNA, 283 Kb region containing skin element.
NID: g2627063. >gp:gp|Z99116|BSUB0013_132 Bacillus subtilis
complete genome (section 13 of 21): from 2395261 to 2613730.
NID: g2634723.

gtggaaaagggggagtcaaaaggtgaaaacgctaattgattaaagcaatggggacggtgata
cgttttatcgattgagcatcaacatccggatacattacttcaagaagctgaaataaaaatt
cgtgcttgggaatcacaaatttagtgctaattgatccgaaatcagatttgatgaatgtgaat
cagcatgcaggtatcgaccagtcagggttagttctgagatgtttaacatgatacgtttt
ggttacgaaactacattatcttctaattttaagatgaacattttgatagggccactagtc
aaattatggaaaattgggttttaagatgcattgaaacctaaagaagaggatatacaacgt
gctttattgtgtatggatcctgaaaatcttgttctaattcaaaaacacatgaagtattt
cttacacaatcaggaattggagattgatttaggagctatagttaaaggctattttgctgat
caattacagcaatacttttttagctcatggtgtatcttctggcattatcgatttaggtggt

aatgttttaacaattggtagacaacccgaaacattagaaaaatggcatgtaggtgtacgt
aatccatttcataaggatacactaccactcggttacattaagcgtagagcatcaatcagtt
gtcacatcaggtatctacgaacgctacttcatacaggaaaaatcaattatctcatcatata
ttggattcaacaacaggttatcctgtagataatgatatcgctagcgtgacaatcatatct
5 gatcatgggattgtatggcgaggtatggagtacaattttagtgttttggtcagtcacaaaaa
aatattgaattattaaatctcattgacggtattgaaggcattattgtgacaagagatgga
agcgttttaatgacttcgaaaaatgcaaaagtatttataa

Sequence 3104

10 VERGSQKVKTLMIKAMGTVIRLSIEHQHPDTLLQEAEIKIRAWESQFSANDPKSDLMNVN
QHAGIAPVKVSSEMFNMIRFGYETTLSSNFKMNILIGPLVKLWKIGFKDALPKKEEDIQR
ALLCMDPENLVLNSKTHEVFLTQSGMEIDLGAIVKGYFADQLQQYFLAHGVSSGIIDLGG
NVLTIQRQPETLEKWHVGVNRNPFHKDTLPLVTLSEVHQSVVTSGIYERYFIQENQLFHII
LDSTTGYPVDNDIASVTIISDHGIDGEVWSTICSFGQSQKNIELNLIDGIEGIIIVTRDG
15 SVLMTSKMQKYL*

Sequence 3105

Contig_0625_pos_1591_4674

is similar to (with p-value 8.0e-25)

20 >sp:sp|P33944|YOJL_ECOLI HYPOTHETICAL 38.5 KD LIPOPROTEIN I
N ADA-OMPC INTERGENIC REGION PRECURSOR. >gp:gp|AE000310|AE00
0310_6 Escherichia coli K-12 MG1655 section 200 of 400 of th
e complete genome. NID: g2367131.
atgaagagtattttatttttaagtaataatgtgaaaatattcacaaaaaaactaggagga
25 tttgctatgagtaaggaaatattcgatacttttaaattttaaatgtggtgccgaattaaaa
aatagagtattaatggcaccatgactatccaagctgggtattttgatggaagtgttaca
tcagaaatgattgattattatcaatttagagctgggtgatgcttcagcaatcattgttgaa
agttgttttgttgaaaatcacggacgaggatttccgggagctatagggtattgataatgat
gacaaaatacctggactcaaacgttttagcagaagcgattcaagctaagggatcaaaagcg
30 attttgcaactttatcatgccggaagaatggcaaatcctaaatttaatgaaggagagcag
ccgatatctgcgagccccattgcagcatataagacctgatgctgtaccacctagagaaatg
acacatgctcaaatcaatcagatgattgatgactttggagaggctacacgtcgcgctata
gaagcggggtttgatggtgtcgaaattcatggcgccaacacatacttattacaacaattt
ttctctccacattctaactcgagacaagattcatggggaggcagtcgtgaaaaacgtaca
35 cgattttccaatcgaagttttgacaaaggttcaacacgtcggttgcgtgaaaaagaggcttct
cattttattataggatatcgatttctcactgaagaaattgaagaaccagggcatacgtttt
gaagataccatgtttttactaaatacattagcagaatatgaacctgttacttccatata
tcagcaaacagttatcaacgtacatctattgtgaatcaagaagatacagaacctttaatt
aataagtacatcaaatgcaaagtgcacagttggcaaaaattccattaattgggtgtaggt
40 agtattgcccacgacaagatgcagaacatgcccttgaactaggatatgatcttttaagt
gttgggaaagcctatttagtggaaccacaatggacagataaaaatttcacaaaacgaagaa
gtagaacaatttgcgatatacatgatcagaaagtacttcacataccatccccctttatgg
aaagtaattggactttatgatttttagataaagaagaagagcagtcgtaaatatgaaaaatta
aaagcacttcaaaaataaaaaagttaaatttaacaaagggtacgtatcatgtctatgcaaaa
45 ggtcataatggcaacttacctatgaaagtccaattatcagaagataagattgtaagatc
gaggtagatgatagcggagagctctgaaggcatagcgaaccagtggttgaacgtttacct
caagatattatcaatgggcaaacactgaatgtagatgtcatttcagggtgcgacagtaaca
agtgaaggcatcgtgcaaggatttcagatgcaattgaacaagcaggagaagaccagat
attttacgggcgcgctcctaaaccagtcgttcagtggtctgatgaggttgttgaagagacg
50 actgatgtcgttggtgattggtacaggaggtgcgggactcagtgagctgctacggcatta
gatgaaggaaaaagaagtcacatcatgcttgagaaatttgcagctatagggtggcaatactatc
cgtacagggtggtcaagtcaacgtgctgagcctaaatggcaaaatgcattcccggcactt
gctgggtgaaaaagagacactcatacagttattaaatcatgatgaaaatgatatagatgaa
gcttacattgaagatttcaatactttaaaccgtcaaattaaagactatcttgaaaatagc
55 agtaatgaagatgaatatctttttagttctgtcgaattacatcgtattcaaacatattta
gggtggtaaacgtaaagatcgtaataatgtcgaaatttcagggtgatttatgatttagttaa
acactcacagataaacgttttggaaatcagtatattggttgaaagacaaagggtgtacattt
gatcggttgcgttgcgtatagatgcctgtgggtgctttatggcgctcgtggtcataaaaccaatg
aaagcacaagggttagagtacattgaaaatttaggagactacgttaaacataatcatggt

cgtatTTTTacagaaactactgcagaaaagttaattaaagaaggtaatcaagttgttggt
 attgaagcacgtaaagcaaatggctgctaaagtgaagattcatacacgtcatgggtgtagtg
 ttggctactgggtgcttttgagcgaatacaaaaaatgctacaacaatataatacgtattgg
 gataatattcctgatgacattaagacaacgaattctcctgcaattacaggtgatgggtata
 5 cgttttaggtgtgcaggcaggcgctgacatcgttaggtatgggattctctcaaagtatgccg
 atttctgatccgaagacaggtgcattgtttactggattaattgtaacaccttcaaacttt
 gtcttcgttaataaggaaggacagcggtttgttaacgaatttgaaagtagagatgtatta
 tctaaggcagcattagaacaaaaagacgggtatcttctatattattgcagatgcaaatatt
 aaagcactagctatgaatacaactgaggataaaaattaatcaagaattagaagacggcact
 10 ttagtaaaaagcagataaccttagaagcattagcccaaaaattaaacattgatacaactact
 tttgtgaacacgattgaaagatataataccttcgtagaacaaggacaagatgaggatttc
 aataaaaaatgcattttgatttaaaaaattgaaaaagcaccattctacgcgacaccacgtaaa
 cctgcaatacatcatactatgggtgggtttaaaaataaacacgcatgcacaagttatagat
 gttgaaggtcatatcattgaaggtttatatgcggtggtgaagttgccggtggtattcat
 15 gctggtaaccggttaggcggaatgcactggcagatatttttacttttggtcgcattgcc
 ggtcaaagtgcgtgaacgaaataa

Sequence 3106

MKSILFLSNVVKIFTKKLGGFAMSKEIFDTFKFKCGAELKNRVLMAPMTIQAGYFDGSVT
 20 SEMIDYYQFRAGDASAIIVESCFVENHGRGFPGAIGIDNDDKIPGLKRLAEAIQAKGSKA
 ILQLYHAGRMANKPFNEGEQPIASPIAALRPDAVPPREMTHAQINQMIDDFGEATRRAI
 EAGFDGVEIHGANTYLLQQFFSPHSNRRQDSWGGSSREKRTFPIEVLTQVQHVVAEKEAS
 HFIIGYRFSPEEIEEPGIRFEDTMFLNLTAEYEPDYFHHISANSYQRTSIVNQEDTEPLI
 NKYIKMQSAQLAKIPLIGVGSIAQRQDAEHALELGYDLLSVGKAYLVEPQWTDKISQNEE
 25 VEQFVDIHDQKVLHIPSPPLWKVMDFMILDKEEHRKYELKALQNKVKFNKGTYHVYAK
 GHNGNLPKMQLSSEDKIVSIEVDDSGESEGIANPVFERLPQDIINGQTLNVDVISINGATVT
 SEGIVQGIADAIEQAGEDPDILRARPKPVVQWSDEVVEETTDVVVIGTGGAGLSAAATAL
 DEGKEVIMLEKFAAIGGNTIRTGGQVNAAEPKWQNAFFPALAGEKETLIQLLNHDENDIDE
 AYIEDFNTLKRQIKDYLENSNEDEYLFDSVELHRIQTYLGGKRKDRNNVEISGDYDLVK
 30 TLTDNVLESVYWLKDKGVHFDPSFVDMVPGALWRRGHKPMKAQGLEIENLGDYVKNHNG
 RIFTETTAELIKEGNQVVGIEARKANGAKVKIHRHGVVLTGGFGANTKMLQYNTYW
 DNIPDDIKTTNSPAITGDIRLVQAGADIVGMGFSQMMPISDPKTGALFTGLIVTPSNF
 VFNKKEGQRFVNEFESRDVLSKAALQKDGIFYIIADANIKALAMNTTEDKINQELEDGT
 LVKADTLEALAQLNIDTTTFVNTIERYNTFVEQQQDEDFNKNAPDLKIEKAPFYATPRK
 35 PAIHHTMGGLKINTHAQVIDVEGHIEGLYAAGEVAGGIHAGNRLGGNALADIFTFGRIA
 QSAVTK*

Sequence 3107

Contig_0625_pos_7554_4822

40 is similar to (with p-value 3.0e-18)
 >gp:gp|AF061185|AF061185_1 Phytophthora infestans cyst germ
 ination specific acidic repeat protein precursor (car90) gen
 e, complete cds. NID: g3851513.
 gtggatgaaatcggttcattatgggtggcgaagaaatcaagccaggccataaggatgaattt
 45 gatccaaatgcaccgaaaggttagtcaaacaacgcaaccaggtaagccgggggttaaaaat
 cctgatacaggcgaagtagttactccacctgtggatgatgtgacaaaatattggtccagtt
 gatggagatccgatcacgtcaacggaagaaattccattcgacaagaaacgtgaattcaat
 cctgatttaaaaccagggtgaagagcgtgttaacaaaaagggtgaaccaggaacaaaaaca
 attacaacaccaacaaactaagaacccattacacaggggaaaaagggtggaaccaggaacaaaaaca
 50 acagaaaaaataacaaaaaacaggtagatgaaatcacagaatatgggtggcgaagaaatc
 aagccaggccataaggatgaatttgatccaaatgcaccgaaaggttagccaagaggacgtt
 ccaggtaaaccaggagttaaaaaccctgatacaggcgaagtagtcacaccaccagtggtat
 gatgtgacaaaatattggtccagttgatggagatccgatcacgtcaacggaagaaattcca
 ttcgacaagaaacgtgaattcaatcctgatttaaaaccagggtaaagagcggttaaacag
 55 aaagggtgaaccaggaacaaaaacaattacaacaccaacaaactaagaacccattacaggg
 gaaaaagggtggcgaagggtgaaccaacagaaaaagtaacaaaaaacaccagtagatgaaatc
 acagaatatggtggcgaagaaatcaagccaggccataaggatgaatttgatccaaatgca
 ccgaaaggtagtccaagaggacgttccaggtaaaccaggagttaaaaatcctgatacaggc
 gaagtagttactccaccagtggtgatgtgacaaaatattggtccagttgatggagatccg

attacgtcaacggaagaaattccggtttgataaaaaacgcgaatttgatccaaacttagcg
 ccaggtacagagaaagtcggttcaaaaaggtgaaccaggaacaaaaacaattacaacacca
 acaactaagaacccattacacaggggaaaaaagttggcgaaggtgaaccaacagaaaaagta
 acaaaaacacccagtggtgaaatcggttcattatggtggcgaagaaatcaagccaggccat
 5 aaggatgaatttgatccaaatgcaccgaaaggtagccaagaggacgttccaggtaaacca
 ggagttaaaaaccctgatacaggcgaagtagtcacaccaccagtggtgatgtgacaaaa
 tatggtccagttgatggagatccgatcacgtcaacggaagaaattccggtttgataaaaaa
 cggaatttgatccaaacttagcgccaggtacagagaaagtcggttcaaaaaggtgaacca
 ggaacaaaaacaattacaacaccaacaactaagaacccattacaggggaaaaaagttggc
 10 gaaggtgaaccaacagaaaaagtaacaaaacacccagtagatgaaatcggttcattatggt
 ggccaagaaatcaagccaggccataaggatgaatttgatccaaatgcaccgaaaggtagc
 caagaggacgttccaggtaaaccaggagttaaaaatcctgatacaggcgaagtagtcaca
 ccaccagtggtgatgtgacaaaatattggtccagttgatggagatccgattacgtcaacg
 gaagaaattccattcgacaagaaacgtgaattcaatcctgatttaaaaccagggtgaagag
 15 cgtgttaaacagaaaggtgaaccaggaacaaaaacaattacaacaccaacaactaagaac
 ccattaacaggggaaaaaagttggcgaaggtgaaccaacagaaaaataacaaaacaccca
 gtagatgaaatcaacagaatattggtggcgaagaaatcaagccaggccataaggatgaattt
 gatccgaacgcaccgaaaggtagccaagaggacgttccaggtaaaccaggagttaaaaat
 cctgatacaggcgaagtagtcacaccaccagtggtgatgtgacaaaatattggtccagtt
 20 gatggagatccgattacgtcaacggaagaaattccggtttgataaaaaacgcgaatttgat
 ccaaacttagcgccaggtacagagaaagtcggttcaaaaaggtgaaccaggaacaaaaaca
 attacaacaccaacaactaagaacccattacagggagaaaaagttggcgaaggtgaacca
 acagaaaaaataacaaaacacccagtggtgatgagatcggttcattatggtggcgaagaaatc
 aagacaggccataaggatgaatttgatccgaacgcaccgaaaggtagtcaaaacaacgcaa
 25 ccaggttaagccaggagttaaaaatcctgatacaggcgaagtagtcacaccaccagtggt
 gatgtgacaaaatattggtccagttgatggagatccgattacgtcaacggaagaaattccg
 tttgataaaaaacgcgaatttgatccaaacttagcgccaggtacagagaaagtcggttcaa
 aaaggtgaaccaggaacaaaaacaattacaacgccaacaactaagaacccattacaggg
 gaaaaagttggtgaaggtgaaccaactcctaagacacctgttaagagtgatgtcagactg
 30 accgcaatattaacaatactcatatctacatga

Sequence 3108

VDEIVHYGGEEIKPGHKDEFDPNAPKGSQTTQPGKPGVKNPDTGEVVTPPVDDVTKYGPV
 DGDPISTEEIPFDKKREFNPDLPKGEERVQKQGEPTTKTITPTTKNPLTGEKVGEPEP
 35 TEKITKQPVDEITEYGGEEIKPGHKDEFDPNAPKGSQEDVPGKPGVKNPDTGEVVTPPV
 DVTKYGPVDGDPITSTEEIPFDKKREFNPDLPKGERVKQKGEPTTKTITPTTKNPLTG
 EKVGEPEPTEKVTKQPVDEITEYGGEEIKPGHKDEFDPNAPKGSQEDVPGKPGVKNPDTG
 EVVTPPVDDVTKYGPVDGDPITSTEEIPFDKKREFDPNLAPGTEKVVQKGEPTTKTITP
 TTKNPLTGEKVGEPEPTEKVTKQPVDEIVHYGGEEIKPGHKDEFDPNAPKGSQEDVPGK
 40 GVNPNPDTGEVVTPPVDDVTKYGPVDGDPITSTEEIPFDKKREFDPNLAPGTEKVVQKGE
 GTKTITPTTKNPLTGEKVGEPEPTEKVTKQPVDEIVHYGGEEIKPGHKDEFDPNAPKGS
 QEDVPGKPGVKNPDTGEVVTPPVDDVTKYGPVDGDPITSTEEIPFDKKREFNPDLPKGE
 RVKQKGEPTTKTITPTTKNPLTGEKVGEPEPTEKITKQPVDEITEYGGEEIKPGHKDEF
 45 DPNAPKGSQEDVPGKPGVKNPDTGEVVTPPVDDVTKYGPVDGDPITSTEEIPFDKKREFD
 PNLAPGTEKVVQKGEPTTKTITPTTKNPLTGEKVGEPEPTEKITKQPVDEIVHYGGEEI
 KTGHKDEFDPNAPKGSQTTQPGKPGVKNPDTGEVVTPPVDDVTKYGPVDGDPITSTEEIP
 FDKKREFDPNLAPGTEKVVQKGEPTTKTITPTTKNPLTGEKVGEPEPTLKTTPVKS DVRL
 TAILTILIST*

50 Sequence 3109

Contig_0627_pos_5488_4337

is similar to (with p-value 2.0e-86)

>sp:sp|P77364|YBBZ_ECOLI HYPOTHETICAL 38.7 KD PROTEIN IN GI
 P-FDRA INTERGENIC REGION. >gp:gp|U82664|ECU82664_110 Escheri
 55 chia coli minutes 9 to 11 genomic sequence. NID: g1773084. >
 gp:gp|U89279|ECU89279_5 Escherichia coli glyoxylate induced
 proteins GlxB1, GlxB2, GlxB3, GlxB4, GlxB6, GlxB7 and GlxB8,
 and glycerate kinase GlxB5 genes, complete cds. NID: g27352
 35. >gp:gp|AE000157|AE000157_8 Escherichia coli K-12 MG1655

section 47 of 400 of the complete genome. NID: g1786716.

atgtcaccatataaagttatcattgcacctgattcttttaaagaaagtatgtcggcgaag
 gaagctgcttttagctatttaaagatggattccaagagggttcgattccagtacaatatat
 gacattattcctatggctgatgggtggcgaagggaacaaccgaagtattgaaagaagcctta
 5 aatgctacctcttattgtgtagaagtaaaagatccacttaatagaatatcatggctagt
 tatgcgagaagcgacgaacaccaaacagctatcattgaaatggcagctgcttcaggacta
 gcattattgagtaaagatgaaagagatccatctattacaacttcgtacgggtaccggccaa
 ctcattaatgatgcacttaatcacgatgttaataaaaattattttaggaataggtggaagt
 gccacgaatgatgggtgggtgtaggaatgttaaaggcttttaggtgtctcttttaaagataaa
 10 aacaatcaagagattcgcgatggagggttttagccctatctcaaatagaatacattgatatt
 actcgtataaaacccacgattgaaagatgtgaatattaaagtagcctgtgatgtaactaat
 ccattattaggagacaatggagcaacaatagtttatgggtccacaaaaaggcgctcagcaa
 aagatgataccaaagttggattcagcattacgtcactatcatgataaaaattgaaagagaa
 ttaaatatgaatgtaaaagatatcccggtcgctgggtgctgcaggaggcatgggaactgca
 15 ttaatcggtttctaaacgctaaattacgtcctggaattgatgtagtcttgaagagact
 caatttaacaaaggataaaaagatgcgaatttagttgttactggcgaaggtaaaatggat
 aaacaaacaatcttatggcaaaacacccattggcgtagccaaagttgcaaaatcatatgat
 ataccgctcatgtctatttggtagtttaggaaaagattacgaagcaatttatcaccac
 ggtatcgatagcgtgttttagtatcatggaacgtccatgccaccttgacgaagctttgaaa
 20 gaaggcgacttcatgttaaacatacaacaataaatatcgacgacttttacaagtaaaa
 attgaaaaatga

Sequence 3110

MSPYKVIIAPDSFKESMSAKEAALAIKDGFEVFDSSSTIYDIIIPMADGGEGTTEVLKEAL
 25 NATSYCVEVKDPLNRNIMASYARSDEHQTAIIEMAAASGLALLSKDERDPSITTSYGTGQ
 LINDALNHVDNKKIILGIGGSATNDGGVGMALKLVSFKDKNQEI RDGGLALSQIEYIDI
 TRINPRLKDVNIKVACDVNPLLDNGATIVYGPQKGAQQKMI PKLDSALRHYHDKIERE
 LNMNVKDI PGAGAAGGMGTALIAFLNAKL RPIDV VLEETQFKQRIKDANLVVTGEGKMD
 KQTIYKGTPIGVAKVAKSYDIPVIAICGSLGKDYEAIYHHGIDSVFSIMERPCHLDEALK
 30 EGALHVKHTTINIARLLQVKIEK*

Sequence 3111

Contig_0630_pos_4111_3176

is similar to (with p-value 3.0e-25)

35 >sp:sp|P19452|HUTG_KLEAE FORMIMINOGLUTAMASE (EC 3.5.3.8) (F
 ORMIMINOGLUTAMATE HYDROLASE) (HISTIDINE UTILIZATION PROTEIN
 G) (FRAGMENT). >gp:gp|M34604|KPNHUTC_1 K.aerogenes histidine
 utilization repressor C (hutC) gene, complete cds. NID: g14
 9203.

40 atgtatcaacttgcacaatctaattctatggacaggtcggttagatagtgaactgatcct
 acacaatttagacacttccaaactgttaaattcgggtgatttaagtcaatttagattttcg
 gatgaacacaaaggcggtgggttatttaggatatgcaattgataaaggagtagaattaaac
 aaaggacgtgtaggtgcaaaagaagggtcccaatgccattgaagcgagcttttgctggattg
 ccagatttgaatcaatgtgaagagattatagattatggtaattgtagaacacaatcatgag
 45 ttgctaatagatacacagcggaattcgcagatcttgctgctaagtctatcaaacgacat
 aaacaaacatttttacttgggtggcggtcatgatatagcatatgcacaatatttagctact
 cgtaaagtttatcctgagtcgtcaataggtgtgattaatatagatgcgcactttgacaca
 cgcgatgaggggttattcaacctctgggtactagtttttagacagattctagaagaagatgat
 aatgcagattatttagtggttaggtatatctcaaggtggtaatacacaaagctttatttaat
 50 tatgttaaagaaaaagatatcaatttgttatatgcagatgaattactacatcaggtatct
 ccccccattaaagatatgatagaacgttttatccataatcatgatacgggttatgttcaca
 atttgcatggatgtagtagatagtgcatttgcaccaggagtcagtgcaccagctgtccta
 ggaatatatccacatacagttttcgaaacttgctaaacgggtcattccaagtgaaaaagta
 aaatctataagtatagctgaaatgaatccgacgtatgattcagatcaaagaactgctaaa
 55 ttagttgctaatttagtacatcattgttttaatttaa

Sequence 3112

MYQLAQSNLWTGRLDSETDPTQFRHFQTVKFGDLSQLDFSDEHKGVGLLGYAIDKGVELN
 KGRVGAKEGPNAIKRAFAGLPDLNQCEEIIDYGNVEHNHELLIDTQREFADLAAKSIKRH

KQTFLGSGHDIAYAQYLATRKVYPESSIGVINIDAHFDTRDEGYSTSGTSFRQILEEDD
NADYLVLGISQGGNTQALFNYAKEKDIQFVYADELLHQVSPPIKDMIERFIHNHDTVMFT
ICMDVVDSAPAGVSAAPAVLGIYPHTVFEAKRVIPSEKVKISISIAEMNPTYDSDQRTAK
LVANLVHHCCLI*

5

Sequence 3113

Contig_0634_pos_2695_2021

>gp:gp|U40385|SEU40385_1 Staphylococcus epidermidis plasmid
pSK818 insertion sequence IS257(818B) putative transposase g
ene, complete cds. NID: g1762099. >gp:gp|U40386|SEU40386_1 S
taphylococcus epidermidis plasmid pSK818 insertion sequence I
S257(818C) putative transposase gene, complete cds. NID: g17
62101.

atgaactatttccagatataaacaatttaacaaggatgttatcactgtagcgggttggtac
tatctaagatatgcattgagttatcgtgatatatctgaaatattaagggaacgtggtgta
aacgttcatcattcaacggtctaccggttggttcaagaatatgccccattttatatcaa
atttggaagaaaaagcataaaaaagcttattacaaatggcgattgatgagacgtacatc
aaaataaaaaggaaaatggagttatttatatcgtgccattgatgcagagggacatacata
gatatttgggtgcgtaagcaacgagataatcattcagcatatgcgtttatcaaacgtctc
20 attaaacaatttggtaaaccctcaaaaggtaattacagatcaggcaccttcaacgaaggta
gcaatggcctaagtaattaaagggttttaaaacttaaacctgactgtcattgtacatcgaaa
tatctgaataacctcattgagcaagatcacggtcatattaaagtaagaaagacaaggat
caaagtatcaatacagcaaagaatactttaaaagggtattgaatgtatttacgctctatat
aaaaagaaccgcaggctctcttcagatctacggattttcgccatgccacgaaattagcatc
25 atgctagcaagttaa

Sequence 3114

MNYFRYKQFNKDVTITVAVGYLRYALSRYRDISEILRERGVNVHSTVYRWVQEYAPILYQ
IWKKKKKKKAYYKWRIDETIYIKIGKWSYLYRAIDAEGHTLDIWLKQRDNHSAYAFIKRL
30 IKQFGKPQKVITDQAPSTKVAMAKVIKGFKLKPDCHCTSKYLNLLIEQDHRHIKVRKTRY
QSINTAKNTLKGIECIYALYKKNRRSLQIYGFSPCHEISIMLAS*

Sequence 3115

Contig_0635_pos_3548_4750

35 is similar to (with p-value 1.0e-24)

>gp:gp|U33059|APU33059_5 Actinosynnema pretiosum auranticum
diaminopimelate decarboxylase (lysA), 3-amino-5-hydroxybenz
oic acid synthase, oxidoreductase, phosphatase, and aminodeh
ydroquinase synthase genes, complete cds; transcription acti
vator gene, partial cds; and unknown genes. NID: g3056877.
gtgtatcagtcagaaaatcaatcattactttttattgttatttttaggttcattaacagca
tttgccccattggctattgatattgtttttacctggactacctaatttagtcatgatttt
gatatttctgcactctacaactcagcttactatctccttttttatgattggattagcgta
ggaaattttttggctggcccatatctgatattactggtagaaaaaaaccattaattttc
45 ttcactgattatttttactatttgcgagtttaggtattatattcgtcacaatatatggatt
atgattatttttacgatttattcaaggattaactgggtgggtgcaggtgcagtcattctcaaga
gccattgctagtgtatgtactcaggtaatgcgctaactaaatttttatcattattaatg
cttgtcaatggcatttgcgccaattatcgccacctgcgcttggcggtatcattttaaattat
gggccatggcggaattgtatttgaataactaacaatgtttgggattgtcatgttaatagga
50 actttattttaagttcctgagtcgcttgaaaagagcctaagggaagtagtaacataggt
acgatgctaatttaattcaagaactttttaaaacaccccgttttgtattaccatgttg
atacaagggtgtgtcatttgtattactatttacttatatttctgcactctccttttatagtt
caaacaattttatgggttaacgccattaaacttcagtattatgtttgcttttataggcggt
acactcattattttcaagccaattaaaccggaacttgggtgactatatagatagattactg
55 ttgctgagaatcatgtctactatacaagttattgggtgtataatcgtatcactaacttta
ctcaaccattggacttttttgatactttcttgggttgggtgatttttagtggcaccagtt
acagggttgcacggttaggattttcgatagcaatggatgagagtaaaggggccaaggt
agttcttcaagtttgggtgggttgggttcaaaccttacttgggtggcgctcatctctccactt
gttgggtattaaggagacagtaaatgcgataccttatataatcggttatcgttattacagca

ataattcttatggttttacagttgattaatgtgaagatatttaaaaaagctaaaattcat
tga

Sequence 3116

5 VYQSENQSLLFIVILGSLTAFGPLAIDMFLPGLPNISHDFDISASTTQLTISFFMIGLAL
GNFLAGPISDITGRKKPLIFSLIIFTIASLGIIFVTNIWIMIILRFIQGLTGGAGAVISR
AIASDMYSGNALTKFLSLLMLVNGIAPIIAPALGGIILNYGPWRIVFVILTMFGIVMLIG
TLFKVPESLEKSLRESSNIGTMLINFKELFKTPRFVLPMLIQGVSVLFTYISASPFIV
10 QTIYGLTPLNFSIMFAFIGVTLISSQLTGKLVYIDRLLLLRIMSTIQVIGVIIIVSLTL
LNHWTFWILSCGFVILVAPVTGIATLGFSIAMDESKGAKGSSSSLLGLFQTLGGVISPL
VGIKGDSNAIPYIIIVIVITAILMVLQLINVKIFKKAKIH*

Sequence 3117

Contig_0635_pos_1170_214

15 is similar to (with p-value 3.0e-40)

>sp:sp|P28246|BCR_ECOLI BICYCLOMYCIN RESISTANCE PROTEIN (SU
LFONAMIDE RESISTANCE PROTEIN).

atgttactttttttatgttttctaattgaattattacttattgtttttattatatacgaag
caatcgtttactcttaatttatttagttttatcttatataccatcatgtgtttgttatg
20 atgacttatcatatggttaactgtatcaataccatgatgatgtttatcattgtaattgta
gcaatgatactgttattgattaaatcgtttatatttcaagttgcaaacaggacgtttt
tttattttacaacttagtcatcattttttatactgtggggctatttgctgtgagttgttta
tatataagttactattcccctaattatcattaatagcttagctttatgggcccgtatcatt
gcatttagtacaatttattcatttatcggtacttatcttggtctacagcttttgaaaat
25 catcaatattataaacacgtaaagtttaattatggtgcttgagctggaattttctctgaa
gaagtgtactacgtctcttgctgcaagacttgataaagctttatctgtttatcattcaca
cggactaaacctatcatcattgtaagtggtggccaaggtcctgacgaaccaatttcagaa
gcacttgcatgaaaagataccttatagctcacaacgttccggaaccatataatttatg
gagaatcaatccacgaatacacgaaccaatttcttatactcctaaatctatcattcattcg
30 atgatgcttacttcaagtcagatggttggtgtaacaagtcatttcatgttttaagagcg
cttaaatgtgctaaaaaagctcatctttctttcgatggtattggaagtcgtacaccatac
cactttttggcacaatctatgattatagactttttgggtttaatgtatcaatataaaaca
atacttactatttatcttgcgtatggtgttttggttgcaataactacaaccatataa

35 Sequence 3118

MLLFLCFLIELLLIVLLYTKQSFTLNLFSLFYLYTIIGFVMMTYHMTVSIPIYDMFIIVIV
AMILLIKYRYIFKLQGRFFILQLSHHFYTVGLFAVSCLYISTIPLIINSALWAAII
AFSTIYSFIGYLSWSTAFENHQYKHKVLMVLGAGIFSEEVTTLLAARLDKALSVYHSQ
RTKPIIIVSGGQGPDEPISEALAMKRYLIAHNPENHIFMENQSTNTRTNFLYSKSIHS
40 MMPTSSQMLCVTSQFHVLRALKFAKKAHLSFDGIGSRTPYHFLAQSMIIDFLGLMYQYKT
ILTIYFAMLFWLAILQTI*

Sequence 3119

Contig_0637_pos_1621_2709

45 is similar to (with p-value 9.0e-27)

>gp:gp|D45211|ARGOD_1 Arthrobacter sp. gene for opine dehyd
rogenase, complete cds. NID: g1060847.

gtgaaaagaggtaaaaaaatgaaaatagctattgtaggttcaggtaaatggtgcagtaact
gctgcagtggaattggtagataaaaggtcatgatgtacgattatattgtcgtaacgaatct
50 attagtaaatgttagtgcgccctagaaaaaggtggctttgattttaataatgaggagaa
gagaagtttatagagtttactgatattagtgatgatatggagtgatgttttagatggtgca
gacattgttcaggtaattatctcttcattcattcattgaatattatgctaaagtgtgtca
aaatttgtagcgaacgaccatctcattttctttaacattgctgcttcaatgggttcaata
cgatttatgaatgtattagaagatcggtcatattgatgtccatccacacttgcagaagca
55 aatacatatacatatggtacacgtgttgactttaacaatgctaaagtagatttatcttta
aatgttcgctgggtgttcttttcaacatttgatcgtagtgagttaaatgaaagttaata
aaggtatctaaaaatttacgattatcttgtaaaagaagaagtttacttaaaactaatctt
gaaaatggttaaccagaaagtacatcctggaccaacattattgaacgttggacgtattgat
tattcagaagagttttctttatataaagaaggcataacaaaacatactgtgagattatta

catgctattgagatagaacggtttaaattagggagaaaattagggtttgaattatcgact
 gccaaagaatcacgtattcaaaggggtatttagaacggaaagacgaggatgaaccgtta
 aatcgctcttttaatactagtcctgtgttttctcaaattccaggaccgaatcacgttgaa
 aaccgttattttaactgaagatatcgcatatggattagttattatgggtctagtttaggacgt
 5 gtcattgatgtcccgcacctaataatcgatgctgttattatgatagcttcaactattctt
 gaacgcgatttctttgaagagggcctcactatcgaggaattaggcttagataaattagga
 ttagagtaa

Sequence 3120

10 VKRGKKMKIAIVGSGNGAVTAAVDMVDKGDVRLYCRNESISKFDVALEKGGFDFNNEGE
 EKFIETDISDDMEYVLDGADIVQVIIPSSFIEYYAKVMSKFVTNDHLIFFNIAASMGSI
 RFMNVLEDRHIDVPHFAEANTLTYGTRVDFNNAKVDLSLNVRRVFFSTFDRSELNESYE
 KVSKIYDYLVEESLLKTNLENGNPEVHPGPTLLNVGRIDYSEEFSLYKEGITHKTVRLL
 HAIEIERLNLGRKLGFEELSTAKESRIQRGYLERKDEDEPLNRLFNTPVFSQIPGNHVE
 15 NRYLTEDIAYGLVLWSSLGRVIDVPTPNIDAVIMIASTILERDFEGLTIEELGLDKLG
 LE*

Sequence 3121

Contig_0640_pos_7034_6588

20 is similar to (with p-value 9.0e-16)
 >gp:gp|AB012285|AB012285_1 Photobacterium damsela gene for
 sialyltransferase 0160, complete cds. NID: g2988378.
 atgctaattacgagacaacaacctattgccaaagatttacgtatgatgatggcagcactc
 aaaatctctactgactttgaacgaatgggggataatgctgctagtagtgcgtcatatacgt
 25 ttaagagttaaaataaatgataactatgtgtttacacgtttaaaaaccatgggtaaatta
 gcgatgctcatgttagaagatttaataaacgctattagaaataaagatttaccactgata
 aaagaagtcattgagagagatgaagatattgatgattatacgttaacatcgtaataacc
 agttacttaattgataatgacccattcgtagctgggtcaagcacacttagcagctagacac
 ttagaacgaataggtgatcatataagcaatattgctgaaagtgtttattattatttaaca
 30 ggccaacattttgaaacttttgattaa

Sequence 3122

MLITRQQPIAKDLRMMMAALKISTDFERMGDNAASIAHIRLRVKINDNYVFTRLKTMGKL
 AMLMLEDLNNAIRNKDLPLIKEVIERDEDIDDLVYNTVNTSYLIDNDPFAVAGQAHLAARH
 35 LERIGDHISNIAESVYYLTGQHFETFD*

Sequence 3123

Contig_0642_pos_5096_4452

40 is similar to (with p-value 7.0e-22)
 >pir:pir|S56598|S56598 yjjG protein - Escherichia coli >gp:
 gp|U14003|ECOUW93_286 Escherichia coli K-12 chromosomal regi
 on from 92.8 to 00.1 minutes. NID: g1263172. >gp:gp|AE000507
 |AE000507_13 Escherichia coli K-12 MG1655 section 397 of 400
 of the complete genome. NID: g2367380.
 45 gtggatttttatgacgctgagaaaaagcggttttataattagcgagaaatacaatcat
 cagccaactcaacaggatttcgaacattttaagaaagtgaaccaagcgattgggaagca
 tttcaacaaaataaattgactaaagatgaagtgtgtcacaacgatttattaattatttt
 aatgactatcaaattcatgtaaatggaaaagaagctgatgagtgcttagagctgaatta
 gcaaaggccaggttaaattattgatcatatacattagaagttatacaacaattaaaatta
 50 aatcattctcttatatagtaaccaatgggtgtaacagaaacacagctacgacgaattgct
 cagacacaatttaagaaatattcaagatgtctttatatctgaacaagctggatttcaa
 aagtcgatgacagagttcttcgattttgtgtttgaacatatcgagagaataacaggaat
 caaactctaattgtgggagattctttaacgtctgacatttttaggtggtaaaaatgcta
 atatcaacatgttgggttaatttagacaaaaagaaaccatacgtctattcaaccggat
 55 tatatcattaatgatttatcagaaatgattcgcatgttgagtga

Sequence 3124

VDFYDAEKKAFYNLAQKYNHQPTQQDFEHFKKVNQAHWEAFQQNKLTKDEVLSQRFINYF
 NDYQIHVNGKEADECFAELAKAPVKLFDHTLEVIQQLKLNHSLYIVTNGVTETQLRRIA

QTQFNEIFQDVFISEQAGFQKSMTEFFDFVFEHIGENNRRNQT LIVGDSLTS DILGGKNAN
ISTCWFNIRQKENHTSIQPDYIINDLSE MIRIVE*

Sequence 3125

5 Contig_0645_pos_2339_3055
is similar to (with p-value 1.0e-45)
>gp:gp|AB014075|AB014075_8 Clostridium histolyticum genes f
or hypoxanthine-guanine phosphoribosyl-transferase (HGPRTase
) , GTPase and 12 ORFs, complete and partial cds. NID: g38688
10 63.
atggggacgtaaatggaacaacattaaagagaaaaaagcccaaaaagataaaaatactagt
agaatatatgccaaatttggtaaagaaatatatgtagctgcaaagtctggtgagccta
ccagagtc aaatcaaactttaagatttagtattagaacgtgcaaaaacatattcagta
aatcatattatagatagagctattgataaggctaaaggcgctggtgacgaaaactacg
15 cacttaagatatgaagggttttggccgaatgggttcaatgcttatagttgacgcatta
aacaatgtaaatcggtacagcatcagatgtacgtgctgcttcggtaagaatggaggaa
atgggagtatctgggttcagtagcttatatgtttgaccatactgcaacctttggtgtag
ggtaaatctgtagatgaagctcttagaaacactaatggagcaagatattgatgtaagag
gtaattgatgacaatggcttgactattgtttacgcagaaccagatcaatttgcacaagt
20 caggatgcattacgtgaagctggcgttgaggaatttaaagtagcagagtttgaaatg
cctcaaactgatattgagttgtctgaagaggatcaagctatttttgaaaaattaatcg
gcacttgaagacttgggaagatgttcaaaatgttttcataatgtagatttaaaataa

Sequence 3126

25 MGRKWNNIKEKKAQKDKNTSRIYAKFGKEIYVAAKSGEPNPESNQTLRLVLERAKTYSVP
NHIIDRAIDKAGAGDENYDHLRYEGFGPNGSMLIVDALTNVNRRTASDVRAAFGKNGGN
MGVSGSVAYMFDHTATFGVEGKSVDEVLETLMEQDIDVRDVIDDNGLTIVYAE PDQFAQV
QDALREAGVEEFKVAEFEMLPQTDIELSEEDQAIFELIDALELLEDVQNVFHNVDLK*

30 Sequence 3127

Contig_0645_pos_4722_5774
is similar to (with p-value 1.0e-57)
>sp:sp|P39676|FHP_YEAST FLAVOHEMOPROTEIN (HAEMOGLOBIN-LIKE
PROTEIN) (FLAVOHEMOGLOBIN) .
35 atgtttgaggctaaccagaattattaaacatgtttaaccaaacaaccaaagaaaggt
atgcaatctgctgcattagcacaagcagtagctgcagcaatgaatttaataattta
gggtgaattaaagccagcaattatgcctgtggcacataagcactgtgctttacaagttat
cccgaacattatccaattgtaggtgaaaatttacttgcctgcaattcaagatgtgacagg
ttagaaagtgcagcaccagtaattcaaacatgggcaaaagcgtatggagaaattgcagat
40 gtatttatcaaatagagcaagaaatttacaacatattgttatggaaaggttttaacca
tttaaaatcacaaacattacacaagaacaagtgacatcaaatctttcacagttgaatct
gaagaatatgatttaagtcaattcgaaccaggtcaatacattaccgtagatgtttctagt
gaaaagttaccatatagagctaaacgtcattattcaatcatagatggagatgaaaatcac
ttagtatttgggtgtcaaacgtgatgtgactactgaacatgaaggtgaagttcaacaatt
45 ttacatgatgaaatatcagaaggtgacatgattaatttatctgctcctgtaggtggcttt
tcaatagaaaatactgaaagaccgcaattgtttattggttctggcgtaggatgacacca
ttagtttcaatgtttaaaaaagctgcatcattaaacgttccaactcaaatgattcaagcg
gttgtagacagaagatgaacgaccatttgcataaaaacttgatagcattacagataattat
gagaagcacagctacatttacacgtgaaagataaagaaggttatttagaagctaaagaa
50 ttagaacaattatttaagtgaacagcctgaaatttatatttgggtgtgacgaaattctta
cattcaattataaattctcttaagaattaaattatgatatgaatcatgttcattttgaa
acatttatctcctcggtttaaggttcaagtatatag

Sequence 3128

55 MFEANPELLNMFNQTKKGMQSAALAQA VLAAMNINNLGAIKPAIMPVAHKHCALQVY
PEHYPIVGENLLAAIQDVTGLESDDPVIQTWAKAYGEIADVFIKLEQEIYNHMLWKGFKP
FKITNITQETSDIKSFTVESEFYDLSQFEPGQYITVDVSSEKLPYRAKRHYSIIDGDENH
LVFGVKRDVTTEHEGEVSTILHDEISEGDMINLSAPVGGFSIENTERPQLFIGSGVGMTP
LVSMFKKAASLNVPTQMIQAVVTE DERPFAQKLDSITDNYEQQLHLHVKDKEGYLEAKE

LEQYLSEQPEIYICGGTKFLHSIINSLKELNYDMNHVHFETFIPLRSVQV*

Sequence 3129

Contig_0650_pos_7844_8518

5 is similar to (with p-value 1.0e-98)

>gp:gp|U50335|MSU50335_2 Mycobacterium smegmatis phage resistance (mpr) gene, complete cds. NID: g1477566.

atgaactatattcagatataaacaatttaacaaggatgttatcactgtagcgggttggtac
tatctaagatatgcattgagttatcgtgatataatctgaaatattaagggaacgtgggtga
10 aacgttcattcaacaggtctaccgttgggttcaagaatatgcccaattttatatcaa
at ttggaagaaaaagcataaaaaagcattattacaaatggcgtattgatgagacgtacac
aaaataaaaggaaaaatggagctatttatatcgtgccattgatgcagagggacatacatta
gatatttgggttgcgtaagcaacgagataatcattcagcatatgcgtttattaacgtctc
at taaacaatttggtaaacctcaaaaggtaattacagatcaggcaccttcaacgaaggta
15 gcaatggctaaagtaattaaagcttttaaaacttaaacctgactgccattgtacatcgaaa
tatctgaataacctcattgagcaagatcacgcgtcatattaaagtaagaaagacaaggat
caaagtatcaatacagcaaaagaatacttttaaaagggtattgaatgtatttacgctctatat
aaaaagaaccgcaggtctcttcagatctacggattttcgccatgccacgaaattagcatc
atgctagcaagttaa

20

Sequence 3130

MNYFRYKQFNKDVITVAVGYLYRYALSYPDISEILRERGVNVHSTVYRWVQEYAPILYQ
IWKKKKHKAYYKWRIDETIKIKGWSYLYRAIDAEGHTLDIWLKQORDNHSAYAFIKRL
IKQFGKPQKVITDQAPSTKVAMAKVIKAFKLPDCHCTSKYLNLIQDHRHIKVRKTRY
25 QSINTAKNTLKGIECIYALYKKNRRSLQIYGFSPCHEISIMLAS*

Sequence 3131

Contig_0650_pos_6596_5145

>gp:gp|U40385|SEU40385_1 Staphylococcus epidermidis plasmid
30 pSK818 insertion sequence IS257(818B) putative transposase g
ene, complete cds. NID: g1762099. >gp:gp|U40386|SEU40386_1 S
taphylococcus epidermidis plasmid pSK818 insertion sequence I
S257(818C) putative transposase gene, complete cds. NID: g17
62101.

atggttgaaaaattaaaacatgaatgggtttaaccagccaggtaaaaaatatacttgccggt
atcgtgggttgcccttagctttaatccctgaagctatcgcattttcaattattgctggcgta
gatccaatgggttggtttgtatgcttcatttatcatcgctggttggtactgctattgttggt
ggtagacctgcaatgatcaggtgcaacaggggctgttgcccttattagttacaccactt
gtgaaagattatgggtgtagaataatcttttagctgccacgatattaatgggagtaattcaa
40 ttagtttttaggccttctcaaagtggggcgtttaatgaaatttatacctcattccgtcatg
ataggttttgaatgcattaggtattatgattttcatgtcccaatagaacatatcttc
ggtatttcaatctcaacttatatatatgtaattgtaacattactcattgtatatataatc
cctaaatttttcaaagcaataacctgcaccattaatagctatcatcgatttgactgctctt
tatatgtatacaggttctgacgtgagaactgttggtgatttaggaaatattaagcaagct
45 ttaccgcacttttttaattcctaattgttccctttaatttagaaacacttcaaactcattttt
ccatactcgctatctatggctattgttaggtctagtagaaagtttacttactgctaaaatt
gtagatgatgcaacagacacttatagtagtaaaaacagagaatctcgtaggccaaggcatt
gctaataatgattacaggattattcggtgggtatgggaggttggtgccatgattggacaatct
gtaatacaatgtcaaatcaggtgcaaacagtagattatctactttttctgctgggtgtgtc
50 ttaataatctcatgattattgttcttggaggactgttggttcaaaattccaatgccaatttta
gcaggtattatgggttatgggttcgattggtacatttgattggaattcttttaaatatatt
caaaaagcaccaaaaacagatgcagttgttatgatacttacagtataattgtactgatg
acacataacttagctctcgccgtggctcgtaggtgttattttcagtgctttattctttgct
actaaaatatcaaaagtagaagtaacatctgagaagtttggtaaaactaaccgtttatct
55 tttaaagggtcaaatctttttgtttctattgactctatgatggatcaaattagctttaat
attgaaaatagattatagaattaaactttaataatgctcatttatgggatgattcagca
gtagatgctattgtatacaatggtaaggaagttcgaagaaaaaaataacattggtcatgta
gaaaaactaaattcagatagtcgtaaaatagctctcagaattaagcaaaactaaatgaaat
catttaaaactaa

Sequence 3132

MVEKLKHEWFNQPGKNILAGIVVALALIPEAIAFSIIAGVDPMVGLYASFIIAVVTAIVG
GRPAMISGATGAVALLVTPLVKDYGVEYLLAATILMGVIQLVLGLLKVGRMLKFIPIHVSVM
5 IGFVNALGIMIFMSQIEHIFGISISTYIYVIVTLLIVYIIPKFKAIPAPLIAIIVLTAL
YMYTGSVDVRTVGDGDLGNIKQALPHFLIPNVFPNLETQIIFPYSLSMAIVGLVESLLTAKI
VDDATDTYSSKNRESRGQGIANMITGLFGGMGCCAMIGQSVINVKSGANSRLSTFSAGVV
LIFMIIVLGGLVVQIPMPILAGIMVMVSIGTFDWSFKYIQKAPKTDVVMILTIVIIIVLM
THNLALGVVGVIFISALFFATKISKVEVTSEKFGKTNRLSFKGQIFFVSIDSMMDQISFN
10 IENSIIELNFNNAHLWDDSAVDAIDTMVRKFEEKNNIVHVEKLNSDSRKIVSELSKLNEN
HLN*

Sequence 3133

Contig_0651_pos_3288_1672

15 is similar to (with p-value 5.0e-34)

>sp:P31679|YAAU_ECOLI HYPOTHETICAL METABOLITE TRANSPORT
PROTEIN IN CARB-KEFC INTERGENIC REGION (ORF65/66). >gp:gp|AE
000114|AE000114_12 Escherichia coli K-12 MG1655 section 4 of
400 of the complete genome. NID: g1786217.

20 atgaaaataaggggatgaaaatatggatttcgttaaatcaaaaactgacttatttagactc
atagacaatgaagcgcaaacatcgacatctaagatgggtttattcttaatttaggaact
atatttttagatgcatatgatattactatttttaggtacaatgactgatcaactcactcaa
cagtttcacttatcaccatcaacgctatctatagtaatgacctctttacctattgggtgca
ttatttgggtgcattacttgggtgtacattagcacatcagtttggacgcaagcatatttta
25 tcaattgccttactaactcactgtaacctctcttgggtgcggcactcgaccaaagtga
attattctaataatatgtcgttgtataatgggttttgcatttgggaatggatagtcagtt
gctttcacttttatttgcggaaataagtaatttaagcacaaaggaagaaatgttaactat
tggcaagtcgtttgggtatgttgcaatagttacttctgcttttagtggtcatttgcgttcttt
attgctaggggctgggtgcacatttgtggagatatgcaattggatttgggtgcacttattgct
30 ttgtcttgtacatcttgcgaattaaatatttacacgaaagtcgacatgggtgattaat
cattattcttttagaaaaagcaactgaatttgaagaaaatattatcataaagacatccac
ctagaggggaacgcttgaagatgatttaagttctgatgtgacttcgccacataattcttgg
acagacttatttaaacgagatatataaaaagaattatcctggcgactgcgatttcaaca
ttacaaggatgcagtagtactatgggtgtcgggtttatacatacctattattgcaacttatctt
35 attagtaaggataaaaataggtgtattattaggtactgctatagtcfaatatagcaggtatt
ctaggcgcatattttaggtgctcaattgacttataaattaggtacacgcaagcttacaatg
ataggcttcacacttgtattactttcaatgggtatgtgtaggactcttttatcatcatcta
ccaatgcttcttaacactttccttattggattatttttatttggccattcaggagggtcct
ggtagtcaaggaaaaacaattgggtgccttatcattcccgactcatttacgttcacaagct
40 actggcttctgtagaatctgtgaagtcgtactggtagtatcataggtacttttgtctttcca
atcattcttctgtgcagtaggtctaacgaatactatgttaatcttgtccattgtcccttta
ctcggaattatcataacagtatctataaaaatgggaagctgtcggttaaggaatacattgtt
gaatataacgcgtacttttggcattaaacgatatagaagatcgataattagaaaagaatta
acattagctttttaaagaagcgaagtcacaaactcagtcgaatggaaagacgtatcattcga
45 ttactacttaattgattacaagccaaaggagattgctatgggttttaaatttggaatccaaa
gttgtttataatgcgattcaacgtagtaaatgtaaacttaaaagaagttttgaataa

Sequence 3134

MKIRDENMDFVKSKTDLFRLIDNEAQTSTSKMVLFLILGTIFLDAYDITILGTMTDQLTQ
50 QFHLSPSTLSIVMTSLPIGALFGALLGGTLAHQFGRKHILSIALLLTLVTSLGAALAPNV
IILIICRCIMGFAIGMDSVPAFTFIAEISNLKHKGRNVNYWQVVVYVAIVTSALVVIAFF
MLGAGAHLWRYAIGFGALIAFVLYILRIKYLHESPTWVINHYSLEKATEFVRKYHDKDIH
LEGTLEDDLSSDVTSPHNSWDLFKPRYIKRIILATAISTLQGMQYVGVGLYIPIIATYL
ISKDKIGVLLGTAIVNIAGILGAYLGAQLTYKLGRKLTMIIGFTLVLLSMVCVGLFYHHL
55 PMLLNTFLIGLFLFGHSGGPGTQGKTIGALSFPTHLRSQATGFVESVSRTGSIIGTFVFP
IILAAGVLTNTMLILSIVPLLGIIITVSIKWEAVGKEYIVEYNATLALNDIERSIIRKEL
TLAFKRSEVKLSRMERRIIRLLNDYKPKEIAMVLNLESKVYVNAIQRSKCKLKRSFE*

Sequence 3135

Contig_0652_pos_1335_2843

>sp:sp|069282|MQO_CORGL MALATE:QUINONE OXIDOREDUCTASE (EC 1.1.99.16) (MALATE DEHYDROGENASE (ACCEPTOR)) (MQO).

gtgagcaaaaaaatggctaataaagagtgcaaaaaatggtgttattattggcgctggtgtc
5 ttaagtacgacatttgggttctatgattaaagaattagaacctgattggaacatcaaactc
tatgaacgcttagatcgctccaggtattgaaagttctaacgaaagaaacaatgccggtaca
ggacatgcggcggttatgtgaattgaactatacagtacaacaacctgatgggtcaattgat
atagaaaaagccaaagaaatcaacgaacaattcgagatttcaaaacaattctgggggtcac
ttagtaaaaaagtggttaacatcagtaaccctagagatttcattaatccacttctcactc
10 agtttcgtaagaggttaaaaaataacggttaaatcttataaaaaaccggttacgaagcaatgcg
taacttccctatgttcgataacatcgaatatacagaagatatcgaagaaatgagaaatgg
atgccattaatgatgacagggcggtactggtaacgaaatcatggcggttagtaaaatcgac
gaaggtacagatgttaactacggtgaattaaactcgtaaaatggcaaaaagtattgaaaaa
catccaaatgctgatgttcaatacaaccacgaagtaattaatttcaatcgctcgtaaaagac
15 ggtatttgggaaggttaaagttaaaaaccgtaattctggagacggtgaaactggttctagct
gattatgtatttatcggtgcagggcggtggcgctattccactattacaaaaaactggtatc
ccagaaagtaaacatcttgggtgattccctatcagtggtcgagttcttaatttgtacaaa
cctgatgtaattaatgaacatgacgtcaaagtatatggtaagaaccaccaggcacacct
ccaatgactgtaccacatttagatacacgttatatcgatgggtgaaagaacattattattt
20 ggaccatttgcgaatattggccctaaattcttaagaaacggttcttaacttagacttatc
aaatcagttaaaccttataacatcacacattactagcatctgcagttaaaaaacttacct
ttaatcaaatactctatcgaccaagtattaatgactaaagaaggttgatgaaccatcta
cgacggttctaccctgaagctcgtgacgaagattggcaattatacactgcaggtaaacgt
gttcaagttatcaaagataactaaagaacacggtaaaggattcattcaatttggtacagaa
25 gttgttaactctaaagaccactctgttatcgactattgggtgaatcacctggagcatca
acttcagtatcagtagccctagaagtttttagagaaaaactttgctgagtatgaaaaagat
tggactccaaaattacaaaaaatgatcccatcatatggtaaatctcttatcgatgatgtt
aagttaatgagagcaactcgtaaacaaacatctaagatttagaattaaattattacgaa
tctaaataa

30

Sequence 3136

VSKKMANKESKNVVIIGAGVLSTTFGSMIKELEPDWNILYERLDRPGIESSNERNNAGT
GHAALCELNVTVPQPDGSIDIEKAKEINEQFEISKQFWGHLVKSGNISNPRDFINPLPHI
SFVRGKNNVFLKNRYEAMRNFPMFDNIEYTEDIEEMRKWMLMTGRTGNEIMAASKID
35 EGTDVNYGELTRKMAKSIEKHPNADVQYNHEVINFRNRKDGIEWVKVKNRNSGDVETVLA
DYVFIGAGGGAIPLLQKTGIPESKHLGGFPIISGQFLICTNPVDVINEHDVKVYGKEPPGTP
PMTVPHLDTRYIDGERTLLFGPFANIGPKFLRNGSNLDFKSVKPYNITLLASAVKNLP
LIKYSIDQVLMTEKGCNMHLRTFYPEARDEDWQLYTAGKRVQVIKDTKEHGKGFQFGTE
VVNSKDHVSIALLGESPGASTSVSVALEVLKNAEYKDWTPKLQKMI PSYGKSLIDDV
40 KLMRATRQTSKDLELNYYESK*

Sequence 3137

Contig_0652_pos_7388_7999

is similar to (with p-value 9.0e-32)

45 >sp:sp|P15029|FECDECOLI IRON(III) DICITRATE TRANSPORT SYST
EM PERMEASE PROTEIN FECDECOLI. >pir:pir|S56513|S56513 citrate-dep
endent iron transport protein fecD - Escherichia coli >gp:gp
|U14003|ECOUW93_200 Escherichia coli K-12 chromosomal region
from 92.8 to 00.1 minutes. NID: g1263172. >gp:gp|AE000499|A
50 E000499_8 Escherichia coli K-12 MG1655 section 389 of 400 of
the complete genome. NID: g1790732.

atgatcataatgatatttccatcagcacctctatttgttcttctctttaggttcatttatc
gggtgctttgacaataagattatttcttctcagttcttatttcaaaatttgatgtaaaaggga
tcaaaatttagcattgataggttagcgataggtgcaatttgtacggccattgtccaattc
55 ttgcttatacgttaattcctcttgatgcaaaatcggttattatgggttgactggtagttta
tacgggtcataatatagtcattttttatagttttattaccatgggtttattatcactgtacct
atagttattgttattagggtatcaacttgatatttttaatttaggtgatcatgtagccatt
gcactaggagcagctgttaaaaaatcttaaaaaatgatttacttatttagcagtaattgta
gcaggtgcttccattgcggttagtaggggtattagtttttttaggtcttatagcacctcat

attgcacgtcaacttgtcgccataaaaaatatacatgttataatcatgtcaggtttggta
ggagcaatattattaacttttggatggttttagcaagaggtatacaacctcctcttgat
attcctgtatga

5 Sequence 3138

MIIMIFPSAPLFLVPLGSFIGALTISIILSVLISKFDVKGSKLALIGLAIGAICTAIVQF
LLIRNPLDANNALLWLGTSLYGHNIIVNFYSLLPWFIITVPIVLLGLGYQLDILNLGDHVAI
ALGARVKILKMILLVLAVMLAGASIAVVGGISFLGLIAPHIARQLVGHKNIHVIIIMSGLV
GAILLTFGDGLARGIQPPLDIPV*

10

Sequence 3139

Contig_0653_pos_2824_2252

is similar to (with p-value 6.0e-31)

15 >sp:sp|P14776|DHSS_SYNPI SOLUBLE HYDROGENASE, SMALL SUBUNIT
(EC 1.12.--) (TRITIUM EXCHANGE SUBUNIT). >pir:pir|S06919|H
QYCSS soluble hydrogenase (EC 1.12.--) small chain - Synech
ococcus sp. (PCC 6716) >gp:gp|X16658|SYNSOLHY_1 Synechococcu
s DNA for the small subunit of soluble hydrogenase. NID: g48
053.

20 atgctaccaccaggtctagcatttgttgccttatagcgatagagcaaaaaaacgatttgc
gatgtaaaaaacaccgagatttctatttagattttaaataaacataaaatcacaagagcaa
aattcaacgcccttcccccctaatgttggctctatttagaggaataaatgcttatgtagaa
cttgtaaaaaaagaaggattaaatcacggttatttcacgccattttaaatacgtaatgcc
ttaagagcagcactaaaggcacttgaattagaattattagtaaaagatgatgctcatgcc
25 tcacctactgtgacctcatttgttccaaaaaatcaagaagaacttaatatcattaaaaat
caacttaaatctcaattcaatataactattgctgggggacaaggacacttaaaggacaa
atthttgagaattgggtcacatggggaaaatatctccttttgatatttttagcagtcgtgtct
gcattggaaattattttaacttctaatagaatgtcaattatattggaacagggataact
caatttatggaggttatttagacatgagtcataa

30

Sequence 3140

MLPPGLAFVAYSRAKKRFADVKTTPRFYLDLNKYIKSQEQNSTPFTPNVGLFRGINAYVE
LVKKEGLNHVISRHFKIRNALRAALKALELELVKDDAHASPTVTSFV PKNQEELNIIKN
QLKSQFNITIAAGQGHLKGQILRIGHMGKISPFIDILAVVSALEIILTSNRNVNYIGTGIT
35 QFMEVIRHES*

Sequence 3141

Contig_0657_pos_4668_5093

is similar to (with p-value 2.0e-19)

40 >gp:gp|AF001974|AF001974_1 Thermoanaerobacter ethanolicus p
utative TrkG gene, partial cds, and putative TrkA, xylose is
omerase (xylA) and xylulose kinase (xylB) genes, complete cd
s. NID: g2581794.

45 gtgagaccgactgtaccgaatgctgatacgacttcgaagagtattttgattaaaggtata
ttggaattgattatggtaagtataaacgtaaccataaccgataaatgcaatagaaatgaga
atagttacgaaagacaactgtatgtatctttctgatatttctctattaaatatagagttg
tttttttctttgctgatacgtattaaaaatagcgattgttgcaataacaaaagttgtaacc
tttataaccacctgcagcactcaatgggtgcacctccaataaacatgagagccataagtaat
aaagctgtcggtgttttaattgtttccaacgtcaattgtgttaaactcctgcagtccttgtt
50 gtcactgattggaaaaatgcatttcctattttttcaattaatcccatgtgtaacatagag
ttttga

Sequence 3142

55 VRPTVPNADTTSKSILIKGILELIMVSINVITIPINAIEMRIVTKDNCMYLSDISLLNIEL
FFSLRIVLKIAIIVAITKVVTIPPAALNGAPPINMRAISNKAVGVLMFPTSIVLNPVAVLV
VTDWKNAPFIPFISINPMCNIEF*

Sequence 3143

Contig_0660_pos_3013_3495

is similar to (with p-value 3.0e-18)

>gp:gp|D89592|D89592_3 *Vibrio alginolyticus* rhlE, KtrA and KtrB genes, complete cds. NID: g3927863.

5 atgaatgaaataagtcctacacgacctataaacattaatatcattaatacagccttggt
atgtcgttgacatcgcttggtcacacctaagcttaaaccacatgtaccaaagcagacatt
acttcaaagaaaatttgtaaaaaggacaacttgcttggtcagtgagatataataatc
atgctaataaatgttaattaatgacgccatagtaaacggcaaatgatcttgtacatcc
ataatatgtacttctctgttgaaaattttaataacctgtttgtcaccagtattattaaaa
ttaataacaaacaaaattaaaaatagcaaatgttgttgttctgattccacctcctacagaa
10 ctgggagatgatcctataaacatcaataatcccattacaatatattgttgcgctcgctgaaa
ctggacacatctatcgtttgcaaacctgcacttctggctggtgatgattggaacaatgca
taa

Sequence 3144

15 MNEISPTRPININIINTALVMSLTSLVTPKLKPHVFNADITSKKICKKDNLPSCVADIII
MLINVINDAIVNTANDLCTSIICTSLKILIPVLSPLVLLKLITNKIKIANVVVLIPPPTE
LGDDPININNPTIFVASLKCDTSIVCKPALLVVDWNNNA*

Sequence 3145

20 Contig_0660_pos_988_47

is similar to (with p-value 3.0e-76)

>sp:sp|Q46807|ARCL_ECOLI CARBAMATE KINASE-LIKE PROTEIN 1. >
gp:gp|U28375|ECU28375_24 *Escherichia coli* K-12 genome; approx-
imately 64 to 65 minutes. NID: g887800. >gp:gp|AE000370|AE0
25 00370_9 *Escherichia coli* K-12 MG1655 section 260 of 400 of the
complete genome. NID: g2367170.

gtgagtgaaatggctaaaattgtagtagcttttaggtggaacgctttaggaaaatcacca
caagaacaacttgaattagtaaaaaatacagctaaatcccttagtaggattaattactaaa
ggtcacgaaattgtgattagtcacggtaatggaccacaagtaggaagtattaaccttggt
30 ctgaattatgcagctgaacacgatcaaggtcctgcttttccatttgctgaatgtggcgct
atgagtcagcctacatcggtatcaacttcaagaaagttacaaaatgaacttcattca
atgggcatagataagcaagttgtcacactagttacccaagtagaagttgatgaaggcgat
ccagcttttaataagtcacgaataaacccatcggtctgttctacactaaagaagaagcaaat
cgtattcaacaggaaaaaggttatcaattttagaagatgctggtcgaggttacgcgtcgc
35 gttgtaccatcaccacaaccaatatctattatcgaactggaaagtattaaaactctagta
gaaatgacacactcgctcgtcgcaggtggaggtggtataccagtcattcggaacag
catgaatagctttaaaggatatagatgccgctcatcgataaagacaaaacaagtgcatatta
ggtgctgatattcactgtgatcaactcattattttaacagcgattgattatgtttatc
aactatcactgaccaaacaagaagcacttaaaacaacaatatagatacgcttaaaaca
40 tatattgaagaagaacaatttgcgaaggcagcatgctacctaataatcgaatctgccatc
tcctttattgaaaataatcctaacggtagcgtgctcatcacatcattaaatcaattagat
gcagcactagaaggtaaaattggcacactcattacaaagtaa

Sequence 3146

45 VSEMAKIVVALGGNALGKSPQEQLVKNLAKSLVGLITKGHEIVISHGNGPQVGSINLG
LNYAAEHDQGPAPFFAECGAMSQAYIGYQLQESLQNELHSMGIDKQVVTLVTVQVEVDEGD
PAFNPSKPIGLFYTKKEANRIQQEKGYQFVEDAGRGYRRVVPSPQPIISIIELSIKTLV
ENDTLVIAAGGGGIPVIREQHDSFKGIDAVIDKDKTSALLGADIHCDQLIILTAIDYVYI
NYHTDQQALKTTNIDTLKTYIEEEQFAKGSMLPKIESAISFIENNPNGSVLITSLNQLD
50 AALEGKIGTLITK*

Sequence 3147

Contig_0661_pos_4592_5317

is similar to (with p-value 5.0e-26)

55 >sp:sp|P23553|XYNC_CALSA ACETYL ESTERASE (EC 3.1.-.-). >pir
:pir|B37202|B37202 acetyl esterase (EC 3.1.1.6) (XynC) - *Caldicell*
ocellum saccharolyticum >gp:gp|AF005383|AF005383_9 *Caldicell*
ulosiruptor saccharolyticus putative transport protein (XynG
, putative transport protein (XynH), xylanase (XynF), xylan

ase (XynE), xylanase (XynD), xylanase (XynA), acetylxylosida
se (XynC) and xylanase (XynB) genes, complete cds. NID: g264
5411. >gp:gp|M34459|CDCXYNAB_2 C.saccharolyticum xylanase A
(XynA), beta-xylosidase (XynB) and acetyl esterase (XynC) ge
5 nes, complete cds. NID: g144295.
atgcatcaaacgattaatgtcattcttccagaagataaaagttactttgatacaaatgaa
aatgcgaaaccattaaaaactatgttattgctacatggtttatcaagtgatacttctct
tatatgagatatacagatagatagaacgctatgcgaataccaccaactagcagtggtgatg
cctaagtctgatcatagtttctattcaaatatggcttatggacatagttattatgactat
10 atactagaagtttatgattatgttcatcaaatatttccattgtctaaaaagagagaagat
aattttatagcaggtcactctatgggaggttatggtgcaatcaaatgtgcattaacgcaa
agttatcgtttctcaaaagccgctatgctttcagcgccatatgatgttctatgattggt
caatatcaatggtagattttactccagaagcgattgtaggtaatacgaacatgtcgcg
gggacatcttttgatccatactattttagttgaacaagcaatagacaatggacaaacgta
15 ccacaactatatattacttgtggaactgaagatgaattgtatcaaggtaaatattgatttt
gtgaactattttagatgaaaaaggtatttcatatcaatttaaaaaagcgccaggtcatcac
gattatgcattttgggataaaagcaatagaagatgtcattgaccgttttacatcatcacat
atntaa

20 Sequence 3148

MHQITINVILPEDKSYFDTNENAKPLKTMLLLHGLSSDTSSYMRYSIERYANTHQLAVVM
PNADHSFYNSMAYGHSYYDYILEVYDYVHQIFPLSKKREDNFIAGHSMGGYGAIKFALTO
SYRFSKAAMLSAPYDVSMIGQYQWYDFTPEAIVGNTQHVAGTSFDPYYLVEQAIDNGOTL
PQLYITCGTEDELYQGNIDFVNYLDEKGISYQFKKAPGHHDYAFWDKAIEDVIDRFTSSH
25 I*

Sequence 3149

Contig_0667_pos_3956_3216

is similar to (with p-value 2.0e-20)

30 >gp:gp|Y12813|BPA2INT_1 Bacteriophage A2 rep, xis and int g
enes. NID: g3005824.

atgcaagatttacaagtatttaattttgaagatttaccagtaagaaaaatagaagtagat
ggagaaccatatttttttaggtaaagacgtggcagaaatattaggttacacaagatctgat
aatgcaattagaaatcatgttgatgatgaagataagctgacgcaccaagttagtgcatca
35 ggtcaaaaacgaaacatggtaatcatcaacgaatctggtttatacagcttaattctttgac
gctgctaacaacaaagtaaaaacgaaagattagaagaaagctaaacgtttttaaacgttgg
gtaaccgaagatgttttacccttccattcgttaaaacaggtacttatcaagttcctgataat
ccaatggacgcattgcaacttatgttcgacgcacaaaaacaaacaaagaagaatagca
actgttaaagcagatgttattgatatcaaagaaaatcaaaagctagatgcaggagaatac
40 ggattgataacaaaaacagttcatcaacgcgttgcttatatcagacaaattcacggacta
cctaataataaagaagtttaacaaacctttatatagagatattaacagtaacgtaaatatcg
atggctggtatttaaaacaagaacacaatttaaaacaaaaacatttcgatgacgtaaatgaat
atgatcacaatttggtttccatctcaatcaacaatgtatgtcatcaacaattagaatg
gactttgaaaacgaagtataa

45

Sequence 3150

MQDLQVFNFEEDLPVRKIEVDGEPYFLGKDVAEILGYTRSDNAIRNHVDEEDKLTHQVSAS
GQKRNMIINESGLYSLIFDAKQSKNESIRKKAKRFRWVTEVDLPSIRKTGTQYQVPDN
PMDALQLMFDAQQTKEEITVKADVIDIKENQKLDAGEYGLITKTVHQRVAYIRQIHGL
50 PNNKEVKNPLYRDINSNVNTMAGIKTRTQLKQKHFDVMMITNWFPSQSTMYVIKQLEM
DFENEV*

Sequence 3151

Contig_0670_pos_3378_4094

55 is similar to (with p-value 1.0e-45)

>gp:gp|AB014075|AB014075_8 Clostridium histolyticum genes f
or hypoxanthine-guanine phosphoribosyl-transferase (HGPRTase
, GTPase and 12 ORFs, complete and partial cds. NID: g38688
63.

atgggacgtaaattggaacaacattaaagagaaaaagcccaaaaagataaaaaatactagt
 agaatatatgccaaatttggtaaagaaatatatgtagctgcaaagtctggtgagccta
 ccagagtc aaatcaaactttaagattagtagtattagaacgtgcaaaaacatattcagtagct
 aatcatattatagatagagctattgataaggctaaaggcgctggtgacgaaaactacgat
 5 cacttaagatatgaagggttttggccgaatggttcaatgctttagttgacgcattaaca
 aacaatgtaaatcgtagacatcagatgtacgtgctgctgcttcggtagaatggaggaaat
 atgggagtagctggttcagtagcttataatgtttgaccatactgcaacctttggtgtagaa
 ggtaaatctgttagatgaagtccttagaaacactaatggagcaagatatgtatgtaagagat
 gtaattgtatgacaatggcttgactattgtttacgcagaaccagatcaatttgcacaagtt
 10 cagtagcattacgttgaagctggcggttgaggaaatttaaagtagcagagtttgaatgtta
 cctcaaactgatattgagttgtctgaagaggatcaagctatttttgaaaaattaatcgat
 gcacttgaagacttggagatgttcaaaatgttttccataatgtagatttaaaataa

Sequence 3152

15 MGRKWNNIKEKKAQKDKNTSRIYAKFGKEIYVAAKSGEPNPESNQTLRLVLERAKTYSVP
 NHIIDRAIDKAKGAGDENYDHLRYEGFGPNGSMLIVDALTNVNRRTASDVRAAFGKNGGN
 MGVS GSVA YMF DHTATFGVEGKSVDEVLETLM EQDIDVRDVIDDNGLTIVYAEPDQFAQV
 QDALREAGVEEFKVAEFEMLPQTDIELSEEDQAIFEKLI DALEDLEDVQNVFHNVDLK*

20 Sequence 3153

Contig_0673_pos_2813_0

is similar to (with p-value 2.0e-34)

>gp:gp|X81475|MHLMP_1 M.hominis lmp1 and lmp2 genes. NID: g
 587470.

25 atgactgaagcaacaattcaaaattataacgctaaacgtcaaaaagcagagcaagttata
 caaaatgcaataaaattattgaaaacgctcaacctagtgtacaacaagtgtctgatgag
 aaatctaaggtagagcaagcactcagtgaaatgaacaacgccaaatcagcgcttagagct
 gataaacaagaattacagcaagcatataatcagttgattcaaccaacggatttaataat
 aagaaccagcttctatcactgcgtacaatcaaagatatcaacaatttagtaacgaattg
 30 aacagcactaaaacaaatacagatcgctatttaaagagcaaaaatccaagtgtagctgat
 gtcaacaatgcactaaataaagtaagagaagtacaacaaaaattaaacgaagccagagca
 cttttacaaaataaagaagataatagtgcactagttcgagccaaagaacaacttcaacag
 gcagttgaccaagtccttcaacagaaggtatgacgcaacaaactaaagatgattacaat
 tcaaaacaacaagctgctcaacaagaatatcaaaagcacaacaagttatcgataatggc
 35 gatgcgactacacaacaatttctaacgccaaaacaaatgttgaacgcgctttagaagca
 ttaataatgcaaaaactggtttaagagcagataaagaggaacttcaaatgcatataat
 caattaaactcaaaaatttgatacgcgagcgttaaaacgcctgcaagtatcaggaaatacaat
 gaagctaagtcacgtattcaaaactcaaattgattcagctaaaaataaagcaaacagtatt
 ttaacaaatgacaatcctcaagtagtacaagtgactgctgctttaaacaataaaagct
 40 gttcaacctgaattagataaagcgatagcaatgcttaaaaataaagagaataataatgca
 ttggttcaagcgaaacaacaacttcaacaaattgttaatgaagtagatccaacacaaggc
 atgacaacagatactgctaataactataaatcaaaaaaacgtgaagctgaagatgaaata
 caaaaagctcaacaaatcattacaatggcgatgccactgagcaacaaattactaacgaa
 acaaatagagtagaatcaagcgattaatgcaataaacaagccaaaaacgatttacgtgct
 45 gataagctcatttgaaaatgcttataaccaattaatacaaaatgttgatacaaatggg
 aaaaaacctgctagtattcaacaataccaagctgctcgacaagctattgagacgcaatac
 aataacgctaaatcagaagcacatcaaatctctgaaaatagtaacccttcagttaatgaa
 gtagcacaagcattacaaaagttgaagctgtacaacttaaagttaatgacgcgattcat
 atgcttcaaaaataaagagaataatagtgcacttgtcacagctaaaaatcaacttcagcaa
 50 gcagttaatgatatcaaccattacaacaggtatgactcaagattctattaataactatgta
 gctaagaggaattgaggctcaaagtgctatcagaatgcagaagctgtcatcaacaatggc
 gatgcaactgcaaaaacaatttcagacgagaaatctaaagtgaacaagcactagcacat
 ttgaatgatgctaaacagcaattaaactgcagatactactgaattacaacagcagttcaa
 caattaaacagaagaggcgatacaataataaaaagccaagaagtatcaatgcatataat
 55 aaagcaattcaatcattagaaacacaaattacttctgctaaagataatgccaacgctgtg
 atacaaaacctatacgtactgttcaagaggtaataatgcattacaacaagtaaatcag
 ttgaatcaacaataactgaagcaattaatcaacttcaaccgctatcaataatgatgca
 ttaaaagctgcagaagattaaatttagaaaataaaattaatcaactgtacaaactgatggg
 atgacacaacaatctatagaggcttatcaaaacgctaaacgcgtagcccaaaatgaatct

aacactgcttttagcattaattaataacggcgatgccgatgaacaacaaattacaactgaa
 acagaccgagtgcaatcagcaaaactacaaacttaactcaagcaattaacgggttaacagtt
 aataaagaaccattagaaaccgctaaaacagcggttacaaataaacatcgaccaggtacct
 agtacagatgggtatgactcagcaatctgttgcaattataatcaaaaactacaaatagct
 5 aaaaacgaaattaacacaattaataacggttttagcgacaatctagatgttaatgcaatc
 aaaacgaataaagcagaagcgggaacgaatcagtaacgatttaacacaagctaagaataac
 ttacaagttgatactcaaccttttagaaaaataaaaaagacaacttcaagatgaaattgat
 caaggtactaacacagatggaatgactcaagattcagtggaataattacaatgatagctta
 agtgcagcaattatagaaaaaggcaagtaataaattacttaaacgtaatccgacagta
 10 gaacaagttaaagagcggttctaatagcacaacaagtcatacaagatttacaaaatgct
 cgaacttcacttgttccagacaaaactcaacttcaagaagctaaaaatagattagaaaac
 agtattaaccaacaaacagatactgacggcatgactcaagattcgcttaacaattataat
 gataaattagcaaaagctagacaaaaccttgaaaaaatatctaaagtttaggtggtcaa
 cctactgtagctgaaattagacaaaatacagatgaagcaaatgcacataaacaagcatta
 15 gacactgcacgttctcaacttacattaaatagagagccatatatcaatcatattaataat
 gaaagtcatttaaaataacgcgcaaaaagataattttaagctcaagttaactcagcacct
 aatcataatacttttagaaacgattaaaaataaaggctgatactttaaatcaatctatgaca
 gcattaagtgggaagtattgcagattacgaaaaatcaaaaacaagaataatttttagat
 gcattcaacaataaacgtcaagactatgacaatgcagtcattgcggctaaggtatttta
 20 aaccaactcaaagtcgacaatgagtgctgatgtgattgatcaa

Sequence 3154

MTEATIQNYNAKRQKAEQVIQNANKIIENAQPSVQVQSDEKSKVEQALSELNNAKSALRA
 DKQELQQAYNQLIQPTDLNKKPASITAYNQRYQQFSNELNSTKTNTDRILKEQNPSVAD
 25 VNNALNKVREVQKQLNEARALLQNKEDNSALVRAKEQLQAVDQVPSTEGMTQQTCKDDYN
 SKQQAQQEISKAQQVIDNGDATTQQISNAKTNVERALEALNNAKTGLRADKEELQNAYN
 QLTQNIIDTSGKTPASIRKYNEAKSRIQTQIDSANKKANSILTDNPNQVSQVTAALNKIKA
 VQPELDKAIAMLNKENNNALVQAKQQLQIVNEVDPTQGMTTDTANNYKSKKREAEDEI
 QKAQQIINNGDATEQQITNETNRVNQAINAINKAKNDLRADKSQLENAYNQLIQNVDTNG
 30 KKPASIQQYQAARQAIETQYNNNAKSEAHQILENSNPSVNEVAQALQKVEAVQLKVNDIAH
 MLQNKENSALVTAKNQLQAVNDQPLTTGMTQDSINNYVAKRNEAQSAIRNAEAVINNG
 DATAKISDEKSKVEQALAHLDNAKQQLTADTTELQTAVQQLNRRGDTNNKKPRISINAYN
 KAIQSLETQITSAKDNANAVIQKPIRTVQEVNNALQQVNQLNQQLTEAINQLQPLSNNDAL
 LKAARLNLENKINQTVQTDGMTQQSIEAYQNAKRVAQNESNTALALINNGDADEQQITTE
 35 TDRVNQQTNNLTQAINGLTVNKEPLETAKTALQNNIDQVPSTDGMTQQSVANYNQKLQIA
 KNEINTINNVLANNLDVNAIKTNKAEERISNDLTQAKNNLQVDTQPLEKIKRQLQDEID
 QGTNTDGMTQDSVDNYNDSLSAAIIEKGKVNKLLKRNPTVEQVKESVANAQQVIQDLQNA
 RTSVLPDKTQLQEAKNRLENSINQQTDTDGMTQDSLNNYNDKLAKARONLEKISKVLGGQ
 PTVAEIRQNTDEANAHKQALDARSQTLNREPYNHINNESHNNNAQKDNFKAQVNSAP
 40 NHNTLETIKNKADTLNQSMTALSGSIADYENQKQENYLDASNNKRQDYDNAVNAAKGIL
 NQTQSPTMSADVIDQ

Sequence 3155

Contig_0681_pos_7383_6148

45 is similar to (with p-value 4.0e-56)
 >sp:sp|P71359|RECQ_HAEIN ATP-DEPENDENT DNA HELICASE RECQ (E
 C 3.6.1.-). >gp:gp|U32756|U32756_4 Haemophilus influenzae Rd
 section 71 of 163 of the complete genome. NID: g1573729.
 atgtctaagttatcaattggacaaaatgatgttgtcaaaacaagtactaaaagacgcaat
 50 ttaattcttcaaagtcaatccgacttatcagcgacaaaatttgttggattatgttgca
 aatcatgaaggacaggcaggaatcattttattgttccactcgttaagcaggtagaagaatta
 cacgaagctctaaatagtgaaaaaattaagagcacaatttatcatgctggtttaacgaat
 aaagagagaattgagggcgaatgatttcttgtatgatcgtgtagaggttgctcattgcg
 acaaatgcatttggtatgggtattgataaatcaaatgtacgttatgtcattcactataac
 55 atgcctggagatttggaatcttactatcaggaagctggacgcgaggacgtgatggttta
 aaaagtgaagtgtatccttttgttttagtgaacgagataagggattacatgagtattttatt
 actgtatcccaagctgatgatgactataaagataaaaatgggcaaaaataaacgaaaatg
 attcaatataccaaaaacgaaaagtggttagaagcgacaattgttcattattttgaaccc
 aatgaaaatttagaggaatgcaatcaatgtagtaattgtatacaggaaaataaaacgtat

gatatgactcgtgaagcgaaaatgattattagctgtattgctcgaatgaagcaacaggaa
 aattatagtgttattatacaagttttacgtggagaagtgcagattatataaaacaccat
 cattataatgaattaacgacacatgggttgatgaaaaattatacaacatctgagttatca
 cacttaattgatgagctacgtttcaaaggatattttaaatagaaaatgatgaaattcttatg
 5 tgtgatacatcagtgaaacaattactaaataatcataccaaggtttataccactccgttc
 aaacaaaaaactaaagagaaggtatttatcaacactgttgaaggtgtggatagagcgta
 tatcgtgagcttggtgatgtacgtaaacagctaagtataaacttggatagcacctgta
 agtatattttctgattacacgctcgaagaatttgctaagcgtaaactgaatcgaaacaa
 gaaatgattgctattgatggtgtaggttagttataaataaagcattattgtcctaagttt
 10 atcgaaaccatacaaaagctataaaactagaatataa

Sequence 3156

MSKLSIGQNDVVKTSTKRRNLIFKVNPTYQRQKFVVDYVANHEGQAGIIYCSTRKQVEEL
 HEALNSEKIKSTIYHAGLTNKERIEAQNDFLYDRVEVVIATNAFGMGIDKSNVRYVIHYN
 15 MPGDLESYYQEAGRAGRDGLKSECILLFSEKDLGHEFYFITVSQADDDYKDKMGEKLTGM
 IQYTKTKKCLEATIVHYFEPNENLEECSNCIQENKTYDMTREA KMIISCIARMKQQE
 NYSVIIQVLGEVTDYIKHHHYNELTTHGLMKNYTSELSHLIDELRFKGYLNENDEILM
 CDTSVKQLLNNHTKVYTPFKQKTEKVFINTVEGVDRALYRELVDVRKQLSDKLGIAPV
 SIFSDYTLLEEFKRPESKQEMIAIDGVGSYKLKHYCPKFIETIQSYKTRI*

Sequence 3157

Contig_0683_pos_4043_4831

is similar to (with p-value 1.0e-70)

>sp:sp|P23355|PTFB_XANCP PTS SYSTEM, FRUCTOSE-SPECIFIC IIBC
 25 COMPONENT (EIIBC-FRU) (FRUCTOSE- PERMEASE IIBC COMPONENT) (PHOSPHOTRANSFERASE ENZYME II, BC COMPONENT) (EC 2.7.1.69) (E II-FRU). >pir:pir|B40944|B40944 phosphotransferase system enzyme II (EC 2.7.1.69), fructose-specific - Xanthomonas campestris pv. campestris >gp:gp|M69242|XANFRUKAA_3 X.campestris
 30 1-phosphofructokinase (fruK) and PTS enzyme-II fructose (fru A) genes, complete cds. NID: g155366.

gtggttttgttgagtaactcaaaaaaagtagccgttgtcacaggtgcagcacaaggtatt
 ggcttgaaaattgctgagcgtctatttgaagacggatatagcatcgcgctttagacttt
 aatgaagaggttagctaaagagtcagctgaaaaattatcaaaagaagggaagaggcagtt
 35 gcttttaaagcagacgtttcaaatacgcgatcaagtatatttagtgtttaaatcaagtcggt
 gaacactttggcgatttaaatgtccttagttaataatgctggcttggaccaatgacacca
 attgaatcagtaaacctgaacaatttaatacaagttgtaggtgtaacgtaggtggtgta
 ttctggggtatccaagctgcaattgaacaatttgataaattaggacatggcggtaaaatt
 atcaatgccacatctcaagcaggtggtgaaggtaatgctggcttatctctatatagcagt
 40 actaagtttgctgtagaggattgactcaagtagcagcagtgatttagctgagaaaaat
 attacagtcacgcattcgacctggtattgttgaaacaccaatgatgaaaggtatcgct
 gaaaagcttgctgagggaaaataaccaaccaatggaatgggggttgaaacaatttacagat
 caaattgccttaaaacgcttatctaaacctgaagatgtagctaagtagtaagcttctta
 gcaggtagcgattcagattatattactggccaaacaatcatcggtgacgggtggtatgaga
 45 ttccactaa

Sequence 3158

VVLLSNSKKVAVVTGAAQIGLKIARLFEDGYSIALVDFNEEVAKESAELSKEGQEA
 AFKADVSNRDQVFSVLNQVVEHFGLNLVNNAGLGPMPTPIESVTPEQFNQVVGVMVGGV
 50 FWGIQAAIEQFDKLGHGKKIINATSQAGVEGNAGLSLYSSTKFAVRGLTQVAARDLAEKN
 ITVNAFAPGIVETPMKGIKLAEEENQPMWGWKQFTDQIALKRLSKPEDVANVVSFL
 AGSDSDYITGQTIIVDGMGRFH*

Sequence 3159

Contig_0683_pos_0_1268

is similar to (with p-value 3.0e-73)

>sp:sp|Q48436|BUDC_KLEPN ACETOIN(DIACETYL) REDUCTASE (EC 1.
 1.1.5) (ACETOIN DEHYDROGENASE) (AR). >gp:gp|D86412|D86412_1
 Klebsiella pneumoniae gene for meso-2,3-butanediol dehydroge

nase (D-acetoin forming), complete cds. NID: g1468938.
 gtgctcacttctaaagaaatcaaagaagctgatggaatcatcattgctgccgatagacag
 gtagatttatcaagggtttaatggtaaaccctctcatcaatgaaagtgtacgtgaaggatt
 catagacccaaggaattaatacaacgtgtgattgaccaagatgcacaaatttaccatgat
 5 caaaatatttcttcaaataatgtctagagaccaggaagaatctcataaaagtaatatcaa
 atgggtatatcagcatttaataatgaatgggtgttccctcatgggtccatttatcgctcgttggc
 gggtttactcatagctattgccttaactcttggaggacacaccactccaaaaggattagtt
 atccccgaagattcatttttgaaatctattgaaaatattggtagtttatcggtttaaattc
 atggttcccatccttgctgggttatatcgcggtgagtagttgctgataagcctgggtcttgtt
 10 ccaggttatgattgggtgggtgccattgctgctgatggtagtttatatggaagtgaagcagga
 gccgggttcttgggtgggtatcgctgcaggtttcttagcgggctatatgcaaaatggatt
 aaacagattaaaagttcctaaagctatggctcctattatgcctattattataacctatt
 ctatcttctttaaatagttgggtctcatttttatatttgtaataaggcgccaccaatttcaaat
 atatttgggtgcattaacatcatgggttaaaaggaatgcaagggtgctaacaatcattattctt
 15 gctcttattatttggcgcgatgattgcttttgatatgggaggtccagtaaaacaagtagca
 ttcttattcggttctgcatttaattgctgaaggcaactacactgtgatgggaatgggtgct
 ttagcagtatgtacaccaccgattgggtttaggttttagctacatttgttcgtaaacaccaa
 ttcaataaagcagaacaagaaatgggtaaggcatcatttacgatgggattatttgggtatt
 actgaaggggcaatcccttttgctgcacaagatcctctaagaatcattccagccaatatg
 20 attggcgcgatgattgcttcagtaataagcggcggttggaggtgtcgggtgataaagttgct
 catggaggtcctattgtcgtgtactaggtggaataagtaaatattttatgggtctttata
 gctgttgcgttgggaagcttagtaaacatgttcacagtcttgttattttaagcgtcacacc
 cctgcTTT

25 Sequence 3160
 VLTSKEIKEADGIIIAADRQVDSLRFNGKPLINESVREGIHRPKELIQRVIDQDAQIYHD
 QNISSNMSRDQEEHSKSNIQMVYQHLNMGVSFMPFIVVGGLLIAIALTLGGHTTPKGLV
 IPEDSFWKSIENIGLSLKFVMPILAGYIAVSIADKPGLVPGMIGGAIAADGSLYGSEAG
 AGFLGGIVAGFLAGYIAKWIKQIKVPKAMAPIMPIIIIPILSSLIVGLIFIFVIGAPISN
 30 IFGALTSWLKGMQGANIIILALIIGAMIAFDMGGPVNKVAFVFLGSALIAEGNYTVMGMVA
 VAVCTPPIGLGLATFVRKHQFNKAEQEMGKASFTMGLFGITEGAIPFAAQDPLRIIPANM
 IGAMIASVIAAVGGVGDVAVHGGPIVAVLGGISNILWFFIAVVVGSVLTMTVLLFKRHT
 PAX

35 Sequence 3161
 Contig_0687_pos_5614_4433
 is similar to (with p-value 4.0e-55)
 >sp:sp|Q44681|RISB_BACAM 6,7-DIMETHYL-8-RIBITYLLUMAZINE SYN
 THASE (EC 2.5.1.9) (DMRL SYNTHASE) (LUMAZINE SYNTHASE) (RIBO
 40 FLAVIN SYNTHASE BETA CHAIN). >gp:gp|X95955|BARIBGENS_4 B.amy
 loliquefaciens ribB, ribG, ribA, ribH & ribT genes. NID: g15
 92687.
 atgcaattcgatacaattgagttggctatatagaggctttaagaaatggagagagcattatt
 gtagttgacgatgaagatagagaaaatgaaggagatctttagctgttacggaatggatg
 45 gatgataataaccattaattttatggctaaagagggtcgtgggtctgatttgtgcaccaatt
 gataaatctatagctgaaagattaaaactacaatctatggagcaaaaataaactgatatt
 tatggcacacattttactgtaagcattgatcattataaaaactactacaggaatcagtgca
 catgaacgtacacaaacggctagagcactcatagatgaaaataactaatcctgaagatttt
 catcgtccggggcacttatttccacttatagcaaaagagaatgggtgtgttaacacgtaaat
 50 ggtcatactgaagctgcgctagatttggcacggttaacaggagcaccaaccagctggagta
 atctgcgaaattatgaatgatgatgggacaatggctaagggtgaagatctccagtcattt
 aaagaacgccaccatttataaaatgattactataaaaagtttgggtgcttttcgtaaggct
 gttgaacttaattgttaattcttaaggcaaaaggtcaagatgccaactgatttgggtcatttt
 gatattgatgatttacaacgattatagcgatgaagaaatcgtagctattgttaaagga
 55 gatttaaaaagcaatcctaattgtacgtatgcattctgctgtctgactggggatattttt
 catagtcaaagatgtgattgcggggcacaactgaaagcgtcaatgaaatataattgacgaa
 catgggtggaatgattatttatttaccctcaagaaggttagaggaatagggttaattaataag
 ttgcgcgcctatgagttgatagaaaagggttatgatacagttactgcaaatcttgcctctt
 gggtttgatgaggatttgagagattatcatgttgcagctgaaatattaaagtattttgat

ataagtgaattaacttgctcagcaataatcctaaaaaatttgaaggtttagaagattac
ggcattgagatcgtagatagaattgaacttatcggtccagaaacacaatataaccatagt
tatatggaaactaaaaaaaataaaatgggacatttaatatag

5 Sequence 3162

MQFDTIELAIEALRNGESIIVVDEEDRENEGDLVAVTEWMDDNTINFMAKEGRGLICAPI
DKSIAERLKLQSMEQNNTDIYGTHFTVSIIDHYKTTTGISAHERTQTARALIDENTNPEDF
HRPGHLFPLIAKENGVLTRNGHTEAAVDLARLTGAQPAGVICEIMNDDGTMAKGEDLQSF
KERHHLKMITIKSLVAFRKAVELNVNLKAKVKMPTDFGHFDMYGFTTDDYSDEEIVAIVKG
10 DLKSNPNVRMHSACLTGDIHFSQRCDCAQLEASMKYIDEHGGMIYLPQEGRGIGLINK
LRAYELIEKGYDVTANLALGFDEDLRDYHVAEILKYFDISEINLLSNNPKKFEGLDY
GIEIVDRIELIVPETQYNHSMETKKNKMGHLI*

Sequence 3163

15 Contig_0687_pos_4420_3959

>sp:sp|P51695|GCH2_BACAM GTP CYCLOHYDROLASE II (EC 3.5.4.25
) / 3,4-DIHYDROXY-2-BUTANONE 4- PHOSPHATE SYNTHASE (DHBP SYN
THASE). >gp:gp|X95955|BARIBGENS_3 B.amyloliquefaciens ribB,
ribG, ribA, ribH & ribT genes. NID: g1592687.
20 atgaattttgaaggtaaattagttggtaaggatttaaaattgcgattgttgttagtaga
tttaattgattttattactacacgtctacttgaaggggctaaagatacacttattcgtcat
gaagtagaagatacaaatattgatgtagcttatgtgcctggcgcatcgaattccactc
gttgcaaaaaaattagctcaaaaagggtgaatatgatgctgtgattacattaggtgtgtg
attagaggcgcaacttcacattatgactatgtatgtaataagtagctaaagggtgtttct
25 aaagcaaacgacatttcagatactccgggtgatttttggagttctaacaactgaaagtatt
gaacaagcagttgaaagagctggtactaaagctggaaataaagggttcagaagcagcagtt
agtgcgaatcgaaatgggctaataaagcaaatcaattaa

Sequence 3164

30 MNFEGKLVGKDLKIAIIVSRFNDFITTRLLEGAKDTLIRHEVEDTNIDVAYVPGAFEIPL
VAKKLAQKGEYDAVITLGCVIRGATSHYDYVCNEVAKGVSKANDISDTPVIFGVLTTESI
EQAVERAGTKAGNKGSEAAVSAIEMANLIKQIN*

Sequence 3165

35 Contig_0691_pos_489_154

is similar to (with p-value 2.0e-21)
>pir:pir|A55345|A55345 diamine N-acetyltransferase (EC 2.3.
1.57) - Escherichia coli >gp:gp|D25276|ECOSN1A_1 Escherichia
coli gene for spermidine acetyltransferase, complete cds. N
40 ID: g517104. >gp:gp|AE000254|AE000254_5 Escherichia coli K-1
2 MG1655 section 144 of 400 of the complete genome. NID: g17
87862.
gtggtagagttattagaaattaactttatacatagaacttggtgaagtgttaattattatc
gatccgcagttatgcaataatgggtacgcgaaaaaagcctttaaaatggctattgactat
45 gcttttttagtattaaatatgaataagggtatacttatatgtggatattaagaatgagaaa
gcagttacatatctatcaaagtaataatttcgaaatagaagggaacgttaaagggaacacttc
tatacaaggggagaatatagagattgctatgtaatgggcttgtaaaaagggaattgggtt
aataagaatgatgatgatttgcctcatataagatga

50 Sequence 3166

VVELLEINFIHRTCEVLIIIDPQYANNGYAKKAFKMAIDYAFVLNMMNKVLYVDIKNEK
AVHIYQSNFEIEGTLKEHFYTRGEYRDCYVMGLLKRNVNKNDDDLSHIR*

Sequence 3167

55 Contig_0692_pos_844_1881

is similar to (with p-value 9.0e-27)
>sp:sp|P45578|YGAG_ECOLI 19.3 KD PROTEIN IN EMRB-GSHA INTER
GENIC REGION.
atgcataataaacaagaatattagattttatagaaaataataaatatgattatgttgaa

ataagtcacgtattcatgaacgccctgaattaggcaatgaagaaatTTTTgcacgaga
 acattaattgaccaattaagagcaaatacgattcgaaatcgaaacggatattgcaggacat
 gcaacaggatttatagcaacgtatgattctgatatgactggaccggttataggatttcta
 gctgaatatgatgctttacctggctcttggtcacgcatgcgggcataataattattggtact
 5 gctagcgtacttgctgcagtagcactaaaagaagtcgctcgatgaaattggtggtaaagta
 gtcgttttgggatgtcctgctgaagaaggtggggaaaaatggctccgcaaaagcttcttat
 gttaaagcaggtgtcattgatgaaattgatgtagcattgatgattcatcctggaaatgaa
 acttatcgtacaattaatacttttagctgtggatgttcttagatattaaattctatggacgt
 agtgcgcatgcatctgaaaatgcagatgaagcattaaacgcttttagatgcaatgatttca
 10 tatattaatggatatagcacagttaaggcaacacattaaaaaaggacaacgagttcacggg
 gttatttttagacgggtggtaaagcggctaataattatacctgattttacacatgcgagattt
 tacactcgagctacttcgaggagagaacttgatgttttaactgaaaaagtaaaccaaatt
 gcaagagggtgctgctattcaaaactgggtgtgattttgaatttggctctatccagaatggt
 gtaaacgaatttatcaaagcacctaacttgatgatttatttgaaaaatatgcaactgaa
 15 ttaggagaagaagtgtatgatgatgattttggctatggatctacagatcacaggtaatgta
 agtcatgttgtaccaacaatacatccacatattaaaattggttctcgaaatcttgtagg
 acatacccaccgcttttag

Sequence 3168

20 MHNQKILDFIENKYDYVEISHRIHERPELGNEEIFASRTLIDQLRANRFEIETDIAGH
 ATGFIATYDSMTGPVIGFLAEYDALPGLGHACGHNIIGTASVLAVALKEVVDEIGGKV
 VVLGCPAEEGGENGSAKASYVKAGVIDEIDVALMIHPGNETYRTINTLAVDVLDIKFYGR
 SAHASENADEALNALDAMISYINGIAQLRQHIKKQQRVHGVILDGGKAANIIPDFTHARF
 YTRATSRRELDVLTEKVNQIARGAAIQTCDFEFGPIQNGVNEFIKAPKLDLDFEKYATE
 25 LGEEVIDDDFGYGSTDTGNVSHVPTIHPHIKIGSRNLVRYTPPL*

Sequence 3169

Contig_0692_pos_609_148

is similar to (with p-value 8.0e-17)

30 >gp:gp|AF006687|AF006687_1 Enterobacter agglomerans indole-
 3-acetyl-L-aspartic acid hydrolase gene, complete cds. NID:
 g2654566.
 atgaatgtagaaagctttaatttagaccataactaaggttggtgcaccttttattcgtcta
 gccgggactatggaaggtcttaattggtgatgtcatacacaaatatgacattcgtttcaaa
 35 cagcccaataaggaacatatggatatgcctggctctacattccttagagcatttaattggca
 gaaaacatttagaaatcatactgataaagtagtagatttaagtcctatgggttgcaaaact
 ggattctatgtttcatttattaatcatgacgactacgatgacgtattaaatattatcgat
 caaacattgcatgatgtgttaaatgctagcgaagtcccagcttgtaattgaggttcaatgt
 gggtgggctgcaagtcattctttagaaggtgctaaaacaattgctcaagcatttttagat
 40 aaaagagagcaatggaatgacatctacggagaaggtaaataa

Sequence 3170

MNVESFNLDHTKVVPFIRLAGTMEGLNGDVIHKYDIRFKQPNKEHMDMPGLHSLEHLMA
 ENIRNHTDKVVDLSPMGCQTGFYVSFINHDDYDDVLNIIDQTLHDVLNASEVPACNEVQC
 45 GWAASHSLEGAKTIAQAFLDKREQWNDIYGEGK*

Sequence 3171

Contig_0693_pos_5566_6633

is similar to (with p-value 9.0e-84)

50 >gp:gp|X94433|BCASPAMIN_1 B.circulans aspartate aminotransf
 erase gene. NID: g1147556.
 atgctaataagataaagcaagatcatttattcagaccatgtatagcgaattaaaatataat
 actaatgaaattgaaaatagaatgaaagagattgagcaagaaattaacttgactggtagt
 tacacacatacttatgaagaattatcttacgggtgcaaaaatggcatggagaaactcaaat
 55 cgttgatttggttagactgttttgaattcttttaaatgttaaagatgcccagatgtatgt
 gacgaaaaagaattttataaaatttatacatacacatattaaagaagctactaacggcgga
 aaaatcaaaccatatattacaatttttagtcctgaagatacacctaaaatttataataat
 cagttgatctggttatgctggttatgaaaatgttggcgatccatctgaaaaaagggttact
 cgtttagctgaacatctaggttggaaggtaaagggttcaaattttgatattttacctctg

atztatcaattgcctaacgacactataaaaaatacacgaacttccaaatgatattgtaaaa
gaagtttctatacatcatgaacactatcccaagctttcaaaattagggttaaaatgggtat
gcggtacattattttcaaatatggatttaaaaaatcggtgggtattacttaccctacagca
ccttttaattggatgggtatattggaaccgaaattgctgtacgtaatttcacagacacctat
5 cgttataatccttttagaaaaagttgcagaagccttttgaatttgatacacttaaaaaataat
tcatttaataaagatcgagcactcgtagagttaaatcatgctgtgtatcattcattttaa
gctgatgggtgtttctattgttgaccacttaactgcagcgaagcaatttgaaatgtttgaa
cgaaatgaacatcaacaaaacagaaatgttactggtaagtggtcttggctggcacctcca
ctttcaccaactttaacttctaactatcatcatggatatgataatacaatgcatcatacg
10 aatttcttctataaaaaagaagaacctatgaagtgcctttccattaa

Sequence 3172

MLIDKARSFIQTMYSSELKYNTNEIENRMKEIEQEINLTGSYTHTYEELSYGAKMAWRNSN
RCIGRLFWSNLNVKDARDVCDEKEFIKFIHTHIKEATNGGKIKPYITIFSPEDTPKIYNN
15 QLIRYAGYENVGDPSEKKVTRLAEHLGWKGKGSNFDILPLIYQLPNDTIKIHLPNDIVK
EVSIIHHEHYPKLSKLGKLYAVPIISNMDLKIGGITYPTAPFNGWYMTETIAVRNFTDTY
RYNLLEKVAEAFEDTLKNNSFNKDRLVELNHAVYHSFKADGVSIVDHLTAAKQFEMFE
RNEHQQRNVTKGWSWLAAPLSPTLTSNYHHGYDNTMHHTNFFYKKEEPMKCPFH*

20 Sequence 3173

Contig_0693_pos_12681_13970

>gp:gp|U13618|SEU13618_2 Staphylococcus epidermidis 9759 he
at shock protein 10 (hsp10) and heat shock protein 60 (hsp60
) genes, complete cds. NID: g535340.
25 atgatgaatccattagcccaaaaattgaatgatgaaataaagcaatcaagtcagaaagta
ttagatatgatgtcacaaattaggtaagatatgttttatccaaaaggaattttatcgcaa
tctgccgaagcaaaacgcacacatatataatgctactattgggtatggcaacccaaaaaagaa
ggtaaaatgtacgcaaattcacttaaccaaattgttaatgacctacaccggatgaaatt
ttcccatatgcacctcctcaagggtgtagaggaattacgtgatttatggcagaaaaaaatg
30 cttaaagaaaatcccgacttaaggtctaaatctatctctcgtcccatcggtacaaatgct
ctcagcgacgggtctttctctagtagctgatttattttagatagacagatgataccgtctta
ttaccgacacacaactggggtaattataaaacttgatttagcacacgctcaggtgctcat
atcaatacgtatttctatttttgatgactcaggtcacttactacatctgaacttgtaaaa
acattaaaagaatataaaaaagacaaagtattatttttaattatcctaataaccca
35 actgggtacacaccaaataaaaaagaagttaatactattgtaaatgcaattgaagaacta
gctaataaagggtactaaagtagtaactgttgcgatgatgcatactatgggttattttat
gaagaagtttaccacagtcgaattttcacggctttaacacaggtgaaatcttctaaccctt
ttaccagtcggtttggatggagctactaaagaattcttctcttgggggtccgagttggc
tttatgacgtttggaattgatcatgaaacgttaaaaaatgacgctagaagctaaagtaaaa
40 ggattaattcgtagcaatatttcaagttctccactaccttctcaaagtgcattcaaacat
gtacttaaatatcatgagcaatttgataaagaatcgatcaaaatatcaatattttaaaa
gaacgctacgaagtaactaaacaagtagtgatgataataaatatgccaaatattggcaa
gcctatgactttaattcaggatactttatgtcattgaaattaaatcaagtcgatccagaa
gaattacgtaaacatttaattaataactattcaattgggtattattgcttttaaatagtaca
45 gatattcgtattgcctttagttgtgtagaaaaagaagatattccttatgtctttgagtcct
attgctaattgcaattgatgatattaataaa

Sequence 3174

MMNPLAQKLNDEIKQSSPEVLDMMSQLGKDMFYPKGILSQSAEAKRTTYNATIGMATKKE
50 LKMYANSLNQMFNDLTPDEIFPYAPPQGVLELRLDWQKKMLKENPDLKSKSISRPIVTNA
LTHGLSLVADLFVDLDDTVLLPTHNWGNVYKLVFSTRHGAIHINTYSIFDDSGHFTTSELVK
TLKEYKKDKVIIILNYPNNPTGYTPNKKEVNTIVNAIEELANKGTQVVTVVDDAYYGLFY
EEVYQQSIFTALTQVKSSNLLPVRLDGATKEFFSWGFRVGFMTFGIDHETLKNLEAKVK
GLIRSNISSSPLPSQSAIKHVLKYHEQFDKEIDQNINILKERYEVTKQVVYDNKYAKYWQ
55 AYDFNSGYFMSLKLNVQDPEELRKHLINNYSIGIIALNSTDIRIAFSCVEKEDIPYVFES
IANAIDDIK*

Sequence 3175

Contig_0693_pos_16689_15784

unknown; although very good hit (p=0) to *B. subtilis* genome

gtgcaggctagtcgaccaattttaattgttgcggatgaagtagaaggcgatgcacttact
 aatattgttttaaacgctatgctggaacatttactgctgtagcagttaaagccccagga
 5 tttggtgatcgacgtaaagcaatgttagaagacctagcaatattaactggtgctcaagtc
 attactgatgatttaggttttagaacttaaagatgcatctcttgatatgctaggtactgct
 aataaagttgaagtgactaaagatcatacaacagtcgtagatggtaattggtgatgaaaat
 aatattgatgctcgtgtaggtcaaattaaagcacaattgaagaaactgattcagagttt
 gataaagaaaaattacaggaacgtttggcaaaactagctggcggcgtagctgttatcaaa
 10 gtaggggctgcaagtgaacagagcttaaagaacgtaaattaagaattgaagacgcatta
 aattcaacacgtgcggcggtggaagaaggatcgttgcgtggtggtactgcttagtc
 aatataatcaaaaaagtaagtgaattaaagcagaaggatggtgaaacgggtgttaat
 atcgtattaaaagcattacaagcacctgttagacaaattgctgaaatgcaggattagag
 ggttcaattattgttgaacgtttaaaacatgctgaagcggcggttgggttcaatgcagca
 15 acaaatgaatgggttaatatgttagaagaaggatagtagatccaactaaagtaactcgt
 tcagcgttacaacatgcagcaagtgtagctgctatgttcttaacaactgaagcagtcgtt
 gctagtattccagagccagaaaaataatgaacaacctggaatgggtggcatgccaggtag
 atgtaa

20 Sequence 3176

VQASRPILIVAEVEGDALTNIIVLNRMRGFTTAVAVKAPGFGDRRKAMLEDLAILTGAQV
 ITDDLGLLELKDASLDMLGTANKVEVTKDHTTVVDGNGDENNIDARVGQIKAQIEETDSEF
 DKEKLQERLAKLAGGVAVIKVGAASETELKERKLRIEDALNSTRAAVEGIVAGGGTALV
 NIYQKVSEIKAEQDVETGVNIVLKALQAPVRQIAENAGLECSI IVERLKHAEAGVGFNAA
 25 TNEWVNMLEEGIVDPTKVTRSAHQHAASVAAMFLTTEAVVASIPEPENNEQPGMGMPGM
 M*

Sequence 3177

Contig_0693_pos_5374_3905

30 is similar to (with p-value 3.0e-77)

>gp:gp|U59924|SSU59924_1 *Sus scrofa* nitric oxide synthase (NOS) mRNA, complete cds. NID: g1762433.

gtgtaccaatataacgacgatagcttaattgttacacaatgatttatatacaaatatgatg
 gctgaaagctactggaatgatggtatccatgaagaatagcagtggttgatttgattttt
 35 cgaaaaatgccatttaattagtggtatgctggtattcaacggattgaaacgcgttgatgaat
 ttcatcgaaaactttgggtttacaaatgaagatcacatatttaaaatcgatagggtat
 gaagaagattttctaaattacctaagaatttgaaatttacagggaatattaaatctatg
 caagaagggtgaaatttggtttggttaattgagccattattaagagttgaagcacctttaatc
 caagcgcaacttattgaaactattttgttaaatatcattaatctccaaacattaattgca
 40 actaaagccagcgaattcgtcaaatagcaactcatgacactttgatggaatttggtaca
 agaagagctcaagagatcgatgctgcactgtggggcgctagagcagcctttattggaggg
 ttgattctacaagtaattgttagagcaggaatacttttaataatcctgtatctggcaca
 catgcacacgcactagtacaaacataggtgatgagtatatagcattcaaaaagtatgct
 gacgcacataaaaaattgtgtgttcttagttgatacttttcatactttaaaatcaggagta
 45 ccaaccgcaattaaagggtgcaaaagagttaggagatactattaattttataggtatcaga
 ttagattctggtgatattgctgacattctaaagaagctcgtagaatgttagatgaggct
 gggtttacagaagctaaaattatcgcatcaaatgatttggtgagcagactattacaagt
 taaaagcacaaggcgctaaagttagcggatggggagtaggtacaaaactgattacagga
 tatgatcaaccagccttaggtgcagtttataaatgggttctattgaaacagatgatggc
 50 acaattgagtgatcgattaaattatcaataatgctgagaaagttactacaccaggcaaa
 aaaaatgtttatcgattatttaataataaaaacaggcaaggctgagggcgactatattacg
 ctagaagggtgaaatcctaattgacgaatctccattgaaatgttccatcctgttcacact
 taaaaatgaagtttattaaatcatttaaagcggttaattctacatcaatctatatgtgaa
 aatggcaacttgatataccatctccagatgaatatgaagctcaggactatcttaaaaat
 55 aatttaagtattttatgggaagaaaaataaacgatatcttaaccgcaagattatccagta
 gatttaagcactaaatgttgggaaaaataagcataagcgtatttttgaagttgctgaacac
 gttaaagagatggaggatgaaaaatgagtag

Sequence 3178

VYQYNDDSLMLHNDLYQINMAESYWNDSIHERIAVFDLYFRKMPFNSGYAVFNGLKRVVN
FIENFGFTNEDITYLKSIGYEEDFLNYLKDCLKFTGNIKSMQEGEICFGNEPLLREAPLI
QAQLIETILLNIINFQTLIATKASRIQIATHDTLMEFGTRRAQEIDAALWGARAFIGG
FDSTSNVRAGKLFNIPVSGTHAHLVQTYGDEYIAFKKYAERHKNCVFLVDTFHTLKSGV
PTAIKVAKELGDTINFIGIRLDSGDIAYLSKEARRMLDEAGFTEAKIIASNDLDEQTITS
5 LKAQGAQVDGWGVGTLITGYDQPALGAVYKLVSIETDDGTMSDRIKLSNNAEKVTPGK
KNVYRIINNKGAEQDYITLEGENPNDESPLKMPHPVHTYKMKFIKSFKAVNLHQSIFE
NGKLVYHLPDEYEAQDYLNKNNLSILWEENKRYLNPQDYPVDLSTKCWENKHKRIFEVAEH
VKEMEDENE*

10

Sequence 3179

Contig_0694_pos_3485_4648

is similar to (with p-value 2.0e-36)

>sp:sp|P31448|YIDK_ECOLI HYPOTHETICAL 62.1 KD PROTEIN IN EM
15 RD-GLVG INTERGENIC REGION. >gp:gp|L10328|ECOUW82_44 E. coli;
the region from 81.5 to 84.5 minutes. NID: g290484. >gp:gp|
AE000445|AE000445_8 Escherichia coli K-12 MG1655 section 335
of 400 of the complete genome. NID: g1790105.

gtgttcaataaaatgtttaaagttgatgaatacttaggtgtaagtagttcaactgctgtc
20 attattatttcatctattatttggtataattggcattatttacttatttataggtggttta
tcgttaagtgcttttagtgattcgatttatggcatggctttaattatagctggacttgcg
attacaatatttaggtctaggtcaattaggagatggcaacttcctacatggtttcgacaaa
atcgtgcaagacacgcctgagaaattgaatggtttggtaaggtggactcggatggtgta
ccttgcccaaccctattcttcggtatgttctttaacaatttattcttctggtgcgcaaac
25 cagatgatagttcaaaaagcactcgcagctaaaaatttaaaagaatctcaaaaagggtgca
atatatttaagcttattttaaagtgttcggacctttatttacagctctaccaggcgtagta
gcatttaactattttaatggtagtcttgataaatcagataacgcttaccctgcacttgta
acttcagtagtaccagaatgggcattcggcttatttgggtgcggttattttgggtgcaata
ttgagctcatttgggtgctcattgaatagtacaactacactattaacactcgatttctat
30 aaacctatttttggaaaaaataaatcagataaacatattgctcgagtaggcatattgct
actgtagtcattggaggttattgtttagcacttgccagtcattctcattattccctagt
ggcttttatgcagtagttcaacagtttaattggtgtgtatagtagccagtgctagcttta
attttaatcgctttcttttctaaacgcacatctaaattggggtgctaaagtgcactcttc
acacatataaattttatagctataatcagctttgtatttacagaaattaattatctatac
35 acatttagtgattattcttttagatttaattattattttgatttttaacaaagttaaa
ctatctagtgagtttgatttaagcacgcaccaaccgaaagtagatatgacgccatggaaa
tatcggtacgttgccaggtattattgttctttagtagtagtaagttatattatcttc
tcaccactcgtgttggtgcaaaaataa

40

Sequence 3180

VFNMFKVDEYLGVSSTAVIIISIIIGIIGIYLFIGGLSLSAFSDSIYGMALIIAGLA
ITILGLGLQDGNFLHGFDKIVQDTPKLNFGFGKVDSDVVPWPTLFFGMFFNNLFFWCAN
QMIVQKALAAKNLKEQKGAIIYLSLFKVFGLFTVLPGVAFNYFNGLSDKSDNAYPALV
TSVLPWEAFGLFGAVIFGAILSSFVGSLSNSTTLLTLDYFKPIFGKNKSDKHIARVGHIA
45 TVVIGVIVVALAPVISLFPVGLYAVVQQFNGVYSMPVLALILIAFFSKRTSKLGAKVTLF
THIILYAIISFVFTEINYLYTFSVLFVFDLIIILIFNKVKLSEFDLSTHQPKVDMTPWK
YRYVAGIIVLALVVVSYIIFSPVLAK*

Sequence 3181

50

Contig_0699_pos_1788_1444

is similar to (with p-value 4.0e-59)

>pir:pir|I67760|I67760 transposase (insertion sequence IS10
) - Escherichia coli >gp:gp|S67119|S67119_2 BST=somatotropin
...BST/beta-Gal fusion protein [Escherichia coli, LBB84, pla
55 smid pXT107, IS10L/R-1, PlasmidTransposonInsertionMutant, 3
genes, 1679 nt]. NID: g455674.

atgcagattgaagaaaccttcgagacttgaaaagtcctgcctacggactaggcctacgc
catagccgaacgagcagctcagagcggttttgatatcatgctgctaatacgccctgatgctt
caactaacatgttggcttgcggcggttcagctcagaaacaagggtgggacaagcacttc

caggctaacacagtcagaaatcgaaacgtactctcaacagttcgcttaggcatggaagtt
 ttgcggcattctggctacacaataacaagggaagactcactcggtggctgcaaccctgctt
 actcaaaatctattcacacatggttacgttttggggaaattatga

5 Sequence 3182

MQIEETFRDLKSPAYGLGLRHSRTSSSERFDIMLLIALMLQLTCWLAGVHAQKQGWDKHF
 QANTVRNRNVLSTVRLGMEVLRHSGYTITREDSLVAATLLTQNLFTHGIVLGKL*

Sequence 3183

10 Contig_0702_pos_12730_11027

>sp:sp|P16659|SYP_ECOLI PROLYL-TRNA SYNTHETASE (EC 6.1.1.15)
) (PROLINE--TRNA LIGASE) (PRORS) (GLOBAL RNA SYNTHESIS FACTO
 R).

atgaaacaatctaaagtattttataccaacgatgagagaagtccttcagaggcggaagca
 15 ttaagccatcgctttatttataaaagcaggggttaattaaacagagtacaagtggtatata
 agttacttaccacttgctacacgtgtactaaataatataatctaaaaatcatagcgaagaa
 atggaaagtatagatgctgtagaaattcttatgccagctttacaacaagcagaattatgg
 gaagagtcaggacgcttgagtgcatatggccagaactaatgcgtttaaagacagaaac
 ggacgtgaatttgcattaggacactcatgaggaagtagtcacttcttagtaagagat
 20 gaattaaaatcatataaacagttaccccttactttatttcaaattcaatccaaatataga
 gatgagaaaagaccacgctttggattattaagaggacgcgaatttcttatgaaagatgca
 tattccttccattcagatgaagcttcattagatgcaacttatcaagatatgtatcaagca
 tatagtcgcatattcaaacgtgtagggcatcaatgcaagaccggttggcagattcaggt
 gcaatagggggaagtcatacacacgaggtttatggcattgagtgaattggggaagataca
 25 atagtttatagtaatgagagtgactatgcagcgaatattgaaaaggctgaagttggttat
 catccttctcataagcattctgcacttgccgaattgactaaagttgagacgccaatggt
 aaacagctcaagaagttgcagaatatttaaagagaccattagatgaaattgtaaaaact
 atgatctttaaataatagatggcgaaatttattatgtttctagttcgtggacatcatgaatta
 aatgaagtgaattaaatcatatttcggcacggaacatgttgaaatggctactccagat
 30 gaaattgtcaatctttagatgccaatccggggtctcttggctctatttttgataaagat
 attaaaatttatgccgataattacttacaagatttaaataactttgttgtaggagcta
 gaagatcattatcactatataaatgtcaatattggttagagactttgatgtaacagaatac
 ggtgactttagatccattacacaaggtgagatgttaagttagctcgggagtagcaca
 tttgctgaaggcattgaagtaggacaagttttcaaatagggacaaaatattctgaatca
 35 atgaatgcaacttttctagataatcagggaaaagctcaaccactcattatgggctgttat
 ggtattggagtatcaagaacatttaagtgcattgttgacaaaaacaatgacgagaatgga
 attatttggccaaaatcagtaacgcctttcgatatccatctaattactatcaatcctaaa
 aaggatgatcaacgtacttttaggtgatcaactttatcaaaaattaatggattcatacgat
 gttttatgatgatgacgaaaagaacgtgctggtgttaatttaattgattcagacctaatt
 40 gggttaccggtacgagttgttggtaaaagagctgaagaaggtattgttgaggtaaaa
 caacgcattaacggtttaagtgaagaagtgcaaatgtatgaattagagattacttaca
 gaattatttaagaatattaagtaa

Sequence 3184

45 MKQSKVFIPTMREVPAAEALSHRLLKAGLIKQSTSGIYSYLPLATRVLNNISKIREE
 MESIDAVEILMPALQQAELWEESEGRWSAYGPELMRLKDRNGREFALGP THEEVVTS
 ELKSYKQLPLTLFQIQSKYRDEKRPRFGLLRGREFLMKDAYSFHSDEASLDATYQDMYQA
 YSRIFKRVGINARPVVADSGAIGGSHTEHFMALSEIGEDTIVYSNESDYAANIEKAEVVY
 HPSHKHSALAEHTKVETPNVKTAQEVAEYLKRPLDEIVKTMIFKIDGFIIMFLVRGHH
 50 NEVKLSYFGTEHVMATPDEIVNLVDANPGSLGPIFDKDIKIYADNYLQDLNMFVVGAN
 EDHYHYINVNIGRDFDVTYGDFFRITQGEMLSDGSGVAQFAEGIEVGQVFKLGTKYSES
 MNATFLDNQKQAQPLIMCYGIGVSRTLSAIVEQNNDENGIIWPKSVTPFDIHLITINPK
 KDDQRTLGDQLYQKLMDSYDVLVDDRKERAGVKFNDSDLIGLPVRVVGKRAEEGIVEVK
 QRINGLSEEVQIDELEYLQELFKNIK*

55

Sequence 3185

Contig_0703_pos_2091_3398

is similar to (with p-value 3.0e-36)

>sp:sp|Q57991|AK_METJA PROBABLE ASPARTOKINASE (EC 2.7.2.4)

(ASPARTATE KINASE). >gp:gp|U67506|U67506_4 Methanococcus jan
naschii section 48 of 150 of the complete genome. NID: g1591
274.

5 gtgaatgaagaccctgaacgtaaaatcattatcgtttcagctccaggcaaaaggcataat
gacgacattaaaactactgattttattaattcgtctctatgaaaaagtacttaataaatta
aattatgaaagtaaaaaacaagaaattatccaaagatatgctgatatagtagaagaatta
ggtataggaaatgacattttaataacaattaatgacacttttagaggaatacattaaacat
ctttctgacaaacctaaccgttttatatgatgctttattatcttgtggcgaaaattttaat
10 gctcaattaatagcccagttataataatagtcagggtattcctactcggttatatttctcct
aaagaagctggattaactgtaactgatttaccacagcaagctcaaatttttagattccgca
tataatgaaatatacaaatgctgatttatgatgaaaagctaattattcctgggtttttc
ggagtttcaaaagcaaaatttatatcggttacgtttccacgcgggtggttctgacataactgg
gctatcatagcagctggcgtccgagcctcactttatgagaacttactgatgtatcagga
atatataaagctaataccgaatatcataaataatcctgaactcatagaggaaataacttat
15 agagaaatgcgagagctatcttatgcaggatttggagtttttcacgatgaagctctacaa
cctttatacaaaagatcgaattcccgtagttatcaaaaataactaatcgctccaaatgataaa
gggacctacattttacatgaccgtgaaatcgattctaaaaatgtcattagtgggaattagt
tgtgataaaggctttactgtgattaatattaaaaaatatttaaatgaatagattagtggga
tttacacgaaagatttcttggcgttttagaagaatttaatatatcatttgaccacatgcct
20 tctgggtattgataacataagtattatcatgctgacaaatcaaattcaaggtaaagaaagt
caagttcttaatgccatacgcaaacggttggaagttgatgaattaagtatcgaccatgat
ttagcagctactaatgattgttgggtgaaggtatgaatcaagttgttggtacagctagtaaa
attactcacgccctttcagaatcaaacattaatttaataatgattaaccaaggtgcttct
gaaatttcaatgatgttttgggaattcatgaagcagatgctgaaaaagcagattatctacg
25 tacgaattttgttacaacggtgtttgtttaaaaaatttgtgtaaataa

Sequence 3186

VNEDPERKIIIVSAPGKRHNDDIKTTDLLIRLYEKVLNKLNYESKKQEIIQRYADIVEEL
GIGNDILITINDTLEEYIKHLSDKPNRLYDALLSCGENFNAQLIAQYNNSSQGIPTRYISP
30 KEAGLTVTDLPQQAQILDSAYNEIYKLRDYDEKLIIPGFFGVSKQNYIVTFPRGSDITG
AIIARGVRASLYENFTDVSIGYKANPNIINNPELIEITYREMRELSYAGFGVFHDEALQ
PLYKDRIPVVIKNTNRPNDKGTIYLHDREIDSKNVISGIDKGFVINIKKYLMLNRLVG
FTRKILGVLEEFNISFDHMPSGIDNISIIMRTNQIQGKESQVLNAIRKRCEVDELSIDHD
LAVLMIVGEGMNQVVGTASKITHALSESNNINLIMINQGASEISMFGIHEADAEKAVLST
35 YEFCYNGVCLKNLCK*

Sequence 3187

Contig_0707_pos_8234_4395

40 >gp:gp|AF067776|AF067776_1 Abiotrophia defectiva extracellu
lar matrix binding protein (emb) gene, partial cds. NID: g32
49002.

gtgaaatctgaagctagacaagcagtagacaataaagcaaatgaacagattaatcatatt
caaaacacgcctgatgcaactaatgaagaaaaacaagaggcaataaatagagtaagtgt
gaattagcaagagttcaagcacaataaatgcagaacatacaaccaagggtgtcaaaact
45 atcaaaagcagcgcgataacttctttatctcgaattaatgcacaagttgttgagaaagag
tctgcaagaaatgcaatcgaaacaaaaggcaacacaaacgcaattttattaataataat
gataatgctacagatgaagaaaaagaggtcgccaacaatttagttatcgctacaaaacaa
aaatcattagataaatattaactccttatcttcaaataatgatgttgaaaatgctaaagta
gcaggaataaatgaaatagctaacgttttaccagcaaccgctgttaagtcaaaagcaaaa
50 aaagatattgatcaaaaactcgcgcaacagattaatcaaattcaaacgcatcaaaactgct
acaactgaggaaaaagaagcggtattcaattggcaaatcaaaaatcaaatgaagcaaga
acagcaattcaaaatgaacatagtaacaatggtgtgcacaagctaataatcaacggcatt
catgaaattgaattagttatgccagatgcgcacaaaaaatctgatgctaaacaaagtatc
gataataaataataatgagcaaaagtaatactatcaacactacaccagatgcaacagatgaa
55 gaaaagcaaaaagcattagataaattaaaaatagctaagatgcaggatacaacaaagtt
gatcaagcgcaaaacaaaccaagaatctgatgcaaaaactgaggctatagatacgata
actaatattcaagcaaatgttgcaaaaaaacatccgctcgagtgggaattagattcaaag
tttgaggatttaaagcgtcaaatcaatgcaacgcccaatgctacagaagaagaaaaacaa
gatgcaattcaaagattgaatggtaaaagagatgaagttaagaatctaataaatcaagat

agacgtgacaatgaagttgaacagcacaaaaatattggacttcaagaattagaacgatt
 catgctaattccaactagaaaatctgatgcgctccaagagttacaaactaaatttatttca
 caaacagagtttaatttaataataacaaaagatgcaactaatgaagaaaaagatgaagccaaa
 cgacttcttgagatttagtaaaaaataaaactataacaaatatcaatcaagcgcaactaat
 5 aatcaagttgataatgctaaagataacggcatgaatgagattgctaccataataccagca
 acaacaattaaaacagatgcaaaaacggctattgataaaaaagctgagcaacaagttaca
 atcatcaatggtaacaacgatgcaacagatgaagaaaaagcagaggctagaaaagctgggt
 gaaaaagcgaaaattgaagccaaatctaattacaaaatagtgatactgaaagggaagtc
 aatgggtgctaaaaaccaaattgggttagaaaaataaacaataattcaaccatcaactcaaact
 10 aaaaacaatgctaagcaagaaaataaattgacaaagctcaagaacaattaatccaaattaat
 aacacgcctgatgcaaccgaagaagaaaagcaagaggcaacaaatagagtcattgctgga
 ttagcacaagcaatacaaaaatttaataatgcacatagtagtcaagaagtaaatgaatct
 aaaacaaatagttattgctacaatcaagagtgtagcaaccaatgtgatcaaaaaacgact
 gctataaatagtttgactcaagaagctaataatcaaaagacggttaatataggtaatgatgg
 15 aatgctactgatgatgaaaaagaggctgcaaagcaattagtgacccaaaaattaaatgaa
 caaattcaaaaaattcatgaaagtacacaagataatcaagttgataacgtaaaagcaca
 gctatcactgcaattaaattgattaatgcaaatgcacataaaagacaagatgccattaat
 attttgactaatctagctgaaagtaaaaaatcagatataagagccaatcaagatgcaact
 actgaagagaaaaatacggcaatacaatctatagatgatacgttagcacaagcacgtaac
 20 aatattaatggtygcaaatacaaatcgcttagtgagatgagaatttagaagatggtaagcaa
 aagttacaacgtattgtgtgtcaactcaaactaaaacacaagctaaagcagacattgct
 caagcaataggtcaacaaaggctgacaatagaccagaatcaaaatgctacaacagaagaa
 aaacaagaagcccttgagagacttaatacaagaacaaatggagtcattgataaatata
 gcagctttagcaaatcaaaatgttacagacgaaaaaaataatatattagaacaataaga
 25 aatgttgaaacctattgtaattgtaaaaccaaaggctaatgaaataattagaaaaaaagct
 gcggaacaaacgactttaataaatcaaaatcaagatgcgacactagaagaaaaacaaata
 gcacttggaataatagaagaagtaagaatgaagcgttaaatcaagtatcacaggcacac
 tcaataatgatgtgaaaattgtggaaaataatggaattgctaaaatttctgaggtccat
 cctgagactataataaacgtaattgctaaacaagaattgaacaagatgcgcaaagtcaa
 30 attgatactatcaatgcaaatataaatacaactaatgaagaaaaatcagccgctatagat
 agagttaatgtagctaaaaattgatgctattaacaatatataatgctacaactacacaa
 ttagttaatgatgctaaaaatagtggttaacacgagtagtagccaaatattaccaagtaca
 gcagtcaaaactaatgcattagcagctctagctagcgaagctaaaaataaaaacgctata
 atagatcaaacaccaaattgcgacagcagaagaaaaagaagaagcaataataaagttgat
 35 cgtcttcaagaagaagcagatgctaatactcctaaaagcacacactactgatgaagttaat
 aatattaaaaatacaagctgttcaaaatattaacgctgttcaagttgaagttatcaagaaa
 caaacgctaaaaaccaaattaaatcaattcattgataatcaaaagaaaattattgaaaat
 acgcctgatgcaacactagaagaaaaagctgaagctaataagattgcttcaaaatgtacta
 acttccacatcagatgaaattgctaattgtagatcataacaacgaggttgatcaagcttta
 40 gataaagctagacccaaaaatcgaggcaattgtaccacaagttagtgaagaaacgagatgct
 ttaaatgcaatccaagaagcatttaattcacaaactcaagaaatacaagagaagcaagaa
 gctacgaatgaagaaaaaactgaagcattaaataaaaataaaccattacttaatcaggct
 aaagtaaatattgatcaagcacagtcacaataaagatgtagatagtgcgaaaacacgtagt
 attcaagatatagagcaaatcaaccacatccacaaacaaaagcaaccgggctcacaga
 45 ttaaatgaaaaagctaaccaacaacaaagtactattgcaactcatcctaattcaacaatt
 gaagaaagacaggaagcaagtgcaaaactacaagaagttcttaaaaaaagccatagctaa

Sequence 3188

VKSEARQAVQNKANEQINHIQNTPDATNEEKQEA INRVSAELARVQAQINA EHTTQGVKT
 50 IKDDAITSLSRINAQVVEKESARNAIEQKATQQTQFINNNDNATDEEKEVANNLVIATKQ
 KSLDNINSLSSNNDVENAKVAGINEIANVLPATAVKSKAKKDIDQKLAQQINQIQTHQTA
 TTEEKEAAIQLANQKSNEARTAIQNEHSNNGVAQAKSNGIHEIELVMPDAHKKSDAKQSI
 DNKYNEQSNTINTTPDATDEEKQKALDKLIAKDAGYNKVDQAQTNQOVSDAKTEAIDTI
 55 TNIQANVAKKPSARVELDSKFEDLKRQINATPNATEEEKQDAIQLNGKRDEVKNLINQD
 RRDNEVEQHKNIGLQLETIHANPTRKSDALQELQTKFISQTELINNNDKATNEEKDEAK
 RLLEISKNTITNINQAQTNQVNDNAKNGMNEIATII PATTIKTDAKTAIDKKAQQVVT
 IINGNNDATDEEKA EARKLVEKAKIEAKSNITNSDTEREVNGAKTNGLEKINNIQPSTQT
 KTNAKQEINDKAQEQLIQINNTPDATEEEKQEATNRVNAGLAQAIQININNAHSTQEVNES
 KTNISATIKSVQPNVIKKPTAINSLTQEANNQKTLIGNDGNATDDEKEAAKQLVTQKLNE

QIQKIHESTQDNQVDNVKAQAITAIKLINANAHRQDAINILTNLAESKKS DIRANQDAT
 TEEKNTAIQSIDDTLAQARNNINGANTNALVDENLEDGKQKLQRIVLSTQTKTQAKADIA
 QAIGQQRSTIDQONQATTEEKQEALERLNQETNGVNDRIQAALANQNVNDEKNNILETIR
 NVEPIVIVKPKANEIIRKKAEEQTLINQNDATLEEKQIALGKLEEVKNEALNQVSQAH
 5 SNNDVKIVENNGIAKISEVHPETIIKRNAKQEIEQDAQSQIDTINANNKSTNEEKSAID
 RVNVAKIDAINNITNATTTQLVNDAKNSGNTSISQILPSTAVKTNALALASEAKNKNAI
 IDQTPNATAEEKEEANNKVDRLQEEADANILKAHTTDEVNNIKNQAVQININAVQVEVIKK
 QNAKNQLNQFIDNQKKIIENTPDATLEEKAEANRLLQNVLTSTSDIEANVDHNNEVDQAL
 DKARPKIEAIVPQVSKKRDALNAIQEAFNSQTQEIQEKQEATNEEKTEALNKINQLLNQA
 10 KVNIDQAQSNKDVDSAKTRSIQDIEQIQPHPQTKATGRHRLNEKANQQQSTIATHPNSTI
 EERQEASAKLQEVLLKSHS*

Sequence 3189

Contig_0711_pos_3905_0

15 is similar to (with p-value 2.0e-53)
 >sp:sp|069282|MQO_CORGL MALATE:QUINONE OXIDOREDUCTASE (EC 1
 .1.99.16) (MALATE DEHYDROGENASE (ACCEPTOR)) (MQO).
 gtgagcaaaaaaatggctaataaagagtcaaaaaatgttggtattattggcgctggtgtc
 ttaagtacgacatttggttctatgattaaagaattagaacctgattggaacatcaaactc
 20 tatgaacgcttagatcgctccaggtattgaaagttctaacgaaagaaacaatgccggtaca
 ggacatgctggcggttatgtgaattgaactatacagtaacaacacctgatgggtcaattgat
 atagaaaaagccaaagaaatcaacgaacaatttcgagatttcaaaacaattctgggggtcac
 ttagtaaaaaagtggttaacatcagtaaccctagagatttcattaatccacttcctcacatt
 agtttcgtaagaggttaaaaaataacgttaaattcttaaaaaaacggttacgaagcaatgcgt
 25 aacttcccctatggttcgataacatcgaatatacagaagatatcgaagaaatgagaaaatgg
 atgccattaatgatgacaggtcgtactggtaacgaaatcatggcggttagtaaaatcgac
 gaaggtacagatgttaactacggtgaattaactcgtaaaatggcaaaaagatttgaaaaa
 catccaaatgctgatgttcaatacaaccacgaagtaattaatttcaatcgctcgtaagac
 ggtatttgggaaggttaaagtta

Sequence 3190

VSKKMANEKSNVVIIGAGVLSTTFGSMIKELEPDWNILYERLDRPGIESSNERNNAGT
 GHAAALCELNYTVQPPDGSIDIEKAKEINEQFEISKQFWGHLVKSGNISNPRDFINPLPHI
 SFVRGKNNVFLKNRYEAMRNFPMFDNIEYTEDIEEMRKWMPLMMTGRGTGNEIMAASKID
 35 EGTDVNYGELTRKMAKSIEKHPNADVQYNHEVINFNRRKDGIEWEVKX

Sequence 3191

Contig_0712_pos_4917_4060

40 is similar to (with p-value 3.0e-91)
 >gp:gp|D13095|BACPK_1 B. stearothermophilus phosphofructoki
 nase and pyruvate kinase genes. NID: g285620.
 atgttttaagatttttttaaatcgaagcaagaaaaagaaatatttaacagttcaagattct
 aaacaaaaatgatgtacctgctggtataatgacaaaatgtcctaattgcaaaaaataatg
 tatacaaaaagaattgaatgaaaatttaaattgtatgctttaattgtgatcatcatatagct
 45 ttaaccgcatataaaaagaatagaagcaatttcagacgatggatcatttatagaatttgat
 agaggtatgacatctgctaacccttagactttcctgggtatgaagaaaaaattgaaaaa
 gatcagcaaaaagactggacttaataagagcgttagtgtctggtagtgcgaaattagatgga
 atacaatatgggtgttgagttatggatgctcgttttagaatgggaagcatgggctctgta
 gttgggtgaaaaaataatgcagaattattgattattgtacagaacatcggttgccatttatt
 50 ctggtttctgcgagtggtggagctagaatgcaagagggaattatttctttaatgcaaatg
 gggaaaaacaagtgtttctttaaaaagacattctgatgcaggactattatatttcttac
 ataactaatcccactactggaggggtttctgcaagttttgcttcggttgagatattaat
 ttaagtgaacctaaagcactaatcggattttgctggtagacgtgttatagaacaaacaatt
 aatgaaaagtgcctgatgatttccaaactgctgagtttttattagagcatgggtcaactt
 55 gataaagtcattcatcgaagatatgcgtgagactttatcaaataattttaaaaatccat
 caagaggtgagtaactaa

Sequence 3192

MFKDFFNRSKKKKYLTVDQSKQNDVPAGIMTKCPNCKKIMYTKELNENLNVCFNCDHHIA

LTAYKRIEAIISDDGSFIEFDRGMTSANPLDFPGYEEKIEKDQKKTGLNEALVSGTAKLDG
IQYGVAVMDARFRMGSMGSMGVGEKICRIIDYCTEHLRPFILFSASGGARMQEGIIISLMQM
GKTSVSLKRHS DAGLLYISYITNPTTGGVSASFASVGDINLSEPKALIGFAGRRVIEQTI
NEKL PDDFQTA EFLLEHGQLDKVIHRKDMRETLSNILKIHQEVSN*

5

Sequence 3193

Contig_0712_pos_3958_3116

is similar to (with p-value 2.0e-74)

10 >sp:sp|Q54776|ACCD_SYN7 ACETYL-COENZYME A CARBOXYLASE CARB
OXYL TRANSFERASE SUBUNIT BETA (EC 6.4.1.2). >gp:gp|U59237|SP
U59237_6 Synechococcus PCC7942 ORF102, ORF120, ORF113, ORF12
8, CTP synthetase (pyrG), carboxyltransferase beta subunit (
accD), ORF145 and ORF123 genes, complete cds. NID: g1399849.

15 atgctggaagcatcattaaaaagagaaactacaaaagtgtacactaatctaaaaccttgg
gatcgtgttcaaatacgctcgtttaccagaaagaccaaccacattagattatattccctat
atTTTTgattcatttattagttacatggcgatagaagttttagggatgatccagcaatg
attgggtggaattgggttacttagatggtaagtctgtaacagttataggccaacaacgtggt
aaagacacgaaagataatatttatcgtaattttggtatggctcaccagaagggtataga
20 aaagctttgctgttaataaagcaagcagagaaatttaatacgtccaataatttacttttata
gatactaaaggtgcttatccgggtaaagcagctgaagaaagaggtcaaagtgaatcaatt
gcaaaaaatttgatggaaatggcttcattaacggtaccagttattgctgttgttattggt
gaaggcgggaagtggcggtcgttttaggaattggaatctcaaatcgtgttctgatgcttgaa
aatagtacttattcagttatttcacctgaaggagcagctgcacttttatggaaagatagt
25 aacttagctcaaattgccgctgaaactatgaaaatcactgcgcgatgatttactagattta
ggtattatagatgaagtgattaatgagccacttgggtggtgcgcaaaaagatgaagaagca
caagctttatcaattaagaaaatgttccttaaacatttaaatgaattaaagcaactcaca
cctgaagaattagcaaatgatcgttttgaaaaatttagaaaaattgggttcagttgtggag
tga

30

Sequence 3194

MLEASLKRETTKVYTNLKPWDRVQIARLPERPTLDYIPYIFDSFIELHGDRSFRDDPAM
IGGIGYLDGKSVTVIGQQRGKDTKDNIYRNFGMAHPEGYRKALRLMKQAEKFNRPITFI
DTKGAYPGKAAEERQSESIKALNMEMASLTVPVIAVVIGEGSGGALGIGISNRVLMLE
35 NSTYSVISPEGAAALLWKDSNLAQIAAETMKITAHDLDDLGIIDEVINEPLGGAQKDEEA
QALSIKKMFLKHLNELKQLTPEELANDRFEKFRKIGSVVE*

Sequence 3195

Contig_0714_pos_982_1668

40 is similar to (with p-value 9.0e-54)

>sp:sp|P54374|AROE_BACSU SHIKIMATE 5-DEHYDROGENASE (EC 1.1.
1.25). >gp:gp|D84432|BACJH642_93 Bacillus subtilis DNA, 283
Kb region containing skin element. NID: g2627063. >gp:gp|Z99
117|BSUB0014_46 Bacillus subtilis complete genome (section 1
4 of 21): from 2599451 to 2812870. NID: g2634966.

45 atgataggaattattggagcaatggaagaagaagtgacgattttaaagcgtaaattgaat
gatatgaatgaaataaatattgcgcatgtttaaattttatgttggcaagctaaaccacaaa
gaggtggttttaacacaaaagtggtataggtaaagttaattgcttctatctcaacgactttg
ttaatagaaaaatttaataccagaagtcgtcattaatactggatcagcaggtgcactagat
50 caaacactatctattggagatatattagtgtgtaatacatgtattatcatcatgatgcta
gctacagcgtttggttatgaatatggacaaataacctcaaatgcctaaaacttatactact
gatcctactttgttgaaaaaaacaatgcatgtattagaacaacaactgaatggtaaa
gtaggtatgattgttagtggtgatagttttataggtagctcagaacagcgacaaaaaatt
aagcaacaatttccagaagctatggctgtcgaaatggaggcaactgcaattgcgcaaaaca
55 tgttatcaatttaaagtaccatttatcgtaactagagctgtttctgatttagcaaacggt
aaagccgatattttctttgaagaatttttagataaagcagctttatcatctagttagagaca
gtttcattattagtagaatcattataa

Sequence 3196

MIGIIGAMEEEVTILKRKLNDMNEINIAHVKFYVGKLNHKEVVLTSQSGIGKVNASISTTL
LIEKFNPEVVINTGSAGALDQTLSDIGDILVSNHVLVHDANATAFGYEGQIPQMPKTYTT
DPTLLKKTMMHVLEQQQLNGKVMIVSGDSFIGSSEQRQKIKQQFPEAMAVEMEATAIAQT
CYQFKVPFIVTRAVSDLANGKADISFEEFLDKAALSSETVSLLES*

5

Sequence 3197

Contig_0714_pos_3323_4141

is similar to (with p-value 1.0e-57)

>sp:sp|P24247|PFS_ECOLI PFS PROTEIN (P46). >pir:pir|S45227|
S45227 purine nucleoside phosphorylase homolog - Escherichia
10 coli >gp:gp|D26562|ECO82K_47 Escherichia coli genome, 2.4-4
.1 min region (110,917-193,643 bp from 0 min). NID: g473770.
>gp:gp|U70214|ECU70214_10 Escherichia coli chromosome minut
es 4-6. NID: g1552727. >gp:gp|AE000125|AE000125_6 Escherichi
15 a coli K-12 MG1655 section 15 of 400 of the complete genome.
NID: g1786348. >gp:gp|U24438|ECU24438_1 Escherichia coli MT
A/SAH nucleosidase gene, complete cds. NID: g2981266.
gtgataaaaagtgaatttgcagtaattggaaacccatttctcattatcgccattg
atgcatcatgctaattttcaatcttttaatttggaaaacacgtatgaagcgataaatgta
20 ccagttaatcaatttcaagacattaaaaaataatttcagaaaagagtattgatggattc
aatgttactattccacataaagaacgtattattccgtacctagatgatattaatgaacaa
gcgaaatctgttggggcggtaaaatacagtttttagttaaagatggtaagtggttggttat
aatactgatgggaattgggtatgtaaatgggtttaaaacaaatatagaaggtatagaagac
gcttatataattaatttttaggtgcaggtggagcaagtaaaaggtatagcaaatgaattatat
25 aaaatcggttcgctccgactttaacagttgcaaatagaacgatgtctcgttttaataattgg
tcgttaaatattaacaaaataaaatttaagccatgcagaaagccatttagatgaatttgat
attataataaacactacacctgctggtatgaacggcaatacagattctgtaatttcttta
aatcgtttagcttcacatacttttagtaagtgatattgtttataatccatataaaacacca
atactaataagaagctgaacaaagaggtaatccaatctataatggtcttgatgatgttcggt
30 catcaaggtgctgaaagttttaaaatttggactaatctagaaccagatataaaagcaatg
aaaaacatagtaattcaaaaattgaaaggagaattatga

Sequence 3198

VIKVKFAVIGNPISHSLSPMHANFQSLNLENTYEAINVVPVNFQDIKKIIEKSIDGF
35 NVTIPHKERIIPYLLDDINEQAKSVGAVNTVLVKDGKWIYNTDGIGYVNLKQIYEGIED
AYILILGAGGASKGIANELYKIVRPTLTVANRTMSRFNNWSLNINKINLSHAESHLDEFD
IIINTTPAGMNGNTDSVISLNLASHTLVSDIVNPNYKTPILIEAEQRGNPIYNGLDMFV
HQGAESFKIWTNLEPDIKAMKNIVIQKLKGEL*

40 Sequence 3199

Contig_0717_pos_7272_6145

is similar to (with p-value 1.0e-99)

>sp:sp|P54524|YQIG_BACSU PROBABLE NADH-DEPENDENT FLAVIN OXI
DOREDUCTASE YQIG (EC 1.-.-.-). >gp:gp|D84432|BACJH642_230 Ba
45 cillus subtilis DNA, 283 Kb region containing skin element.
NID: g2627063. >gp:gp|Z99116|BSUB0013_132 Bacillus subtilis
complete genome (section 13 of 21): from 2395261 to 2613730.
NID: g2634723.
atgaataataaatatgaacctttattttaaatctttaacactacctaattggtggtgaagta
50 agaaatcgttttgttttagctcctcttacacatacttcatcaaatgatgatggaacaatt
tcagatatagaattaccttacattgagaaacgttctaagatggtgggattgcaattaat
gcagcaagtaattgttaattgatgtaggcaagcatttctcggtcaaccttctggtgcacat
gactcagatatcgaagggttaaaagaacttgcacagttatgaagaaaaatggtgcgaaa
gcaatagtcacaaattcatcatggaggtgctcaagctttaccagagttaacacctgatgga
55 gatggtgtgaccaaagtgccatttctcttaaaagttttggtcagcaaaaagaacatgat
gctcgtagatgactgctgaagaaattgaacaaactattagagactttggtgaagcaact
agaagagcaattgaagcaggttttgatggcgttgaaattcatggcgcaaacattatctt
attaccaatttgtttctccttactataatagaagaaatgatgtttgggctgataactat
aaattccctgttgctgttatagatgaagttgttaaagctaaaaaagctcatgcatatgat

gattttattatttggttacagattgtcacctgaagaagcggaatcaccaggtatttcaatg
 gagataactgaagaattaattcaccaaatcgcaaataaaccacttgattatattcatgtg
 tcattaatggatgttaactcagttacgcgagaaggtaaatataaagggtgaaaatcgcttg
 gaacttattcatcaatggataaaatggacgtatgccgcttatttggtataggttctgtcttt
 5 acagctgaagatgcactaaatgctgttgaaaacattggagttgaatttggtgcgttaggt
 tgtgaaattctacttgattatgattttgttgctaaaattaaagaaggctcgagaagacgaa
 attataaatgcttttgatcctaactcgtgaagaccaacattatctaacaccaaattctttgg
 gaacagtttaatacaggattctatccattacctcgaaaagacaaataa

10 Sequence 3200

MNNKYEPLFKSLTLPNGVEVRNRFVLAPLTHTSSNDDGTISDIELPYIEKRSKDVGIIN
 AASNVDVVGKAPFGQPSVAHDSIEGLKELAQVMKKNGAKAIVQIHGGAQALPELTPDG
 DVVAPSAISLKSFGQKQKHEHDAREMTAEIEQTIRDFGEATRRRAIEAGFDGVEIHGANHYL
 IHQFVSPYNNRRNDVWADNYKFPVAVIDEVVKAKKAHAYDDFIIGYRLSPPEEAESPGISM
 15 EITEELIHQIANKPLDYIHVSLMDVNSVTREGKYKGENRLELIHQWINGRMPLIGIGSVF
 TAEDALNAVENIGVEFVALGCEILLDYDFVAKIKEGREDEIINAFDPNREDQHYLTPNLW
 EQFNQGFYPLPRKDK*

Sequence 3201

20 Contig_0718_pos_2272_1247

is similar to (with p-value 3.0e-29)

>sp:sp|P31547|YAE_ECOLI HYPOTHETICAL ABC TRANSPORTER PERME
 ASE PROTEIN YAE. >gp:gp|D83536|ECOTSF_23 Escherichia coli g
 enome, 4.0 - 6.0 min region. NID: g1208942. >gp:gp|U70214|EC
 25 U70214_47 Escherichia coli chromosome minutes 4-6. NID: g155
 2727. >gp:gp|AE000129|AE000129_2 Escherichia coli K-12 MG165
 5 section 19 of 400 of the complete genome. NID: g1786395.
 gtgattgaattcaaaaatgttaacaaagtttttcgcaaaaaaagagaaactattcaagct
 ttgaaaaatgtatcatttaagattgaccaacatgatatttttggtgtatttggttatagt
 30 ggtgctggtaaaaagtcatttagttcggtagtcaatcaacttgagacagtatcagatggt
 caagttattgttgatggtcatgagattgatacatataaagaaaaagatttacgtgatatt
 aaaaaagatatcggtatgatctttcaacatttcaatttgcttaattctaaatcagtcctat
 aaaaatgttgcaatgccacttattttaagtaagacaaataagaaagaaattaaggaaaaa
 gttgacgaaatgtagaatttgtgggcttgctgataaaaaagatcaatttccagatgaa
 35 ttatcaggtggacaaaaacaactggttgccatcgcaagagcattagtaacgcacacctaaa
 atattattatgtgatgaagcgacaagtgtcttgatccagctactacaagctcaatttta
 aatttattaagtaaatgtgaatcgaaacatttggtgtgacgattatgatgattacacatgaa
 atgagcgttaattcaaaaaatttgcacgtgtagctgtcatggaaaatggcgaagtgata
 gaaatggggacagttaaagatgtcttagtcacccacaacgaacactgcaaaaaatttc
 40 gtttcgacggtgattaacactgagccttcaaaagagttacgggcctcttttaactcgaga
 aaagattcaaaatttcacagattataaactgttttagactctgaacaaattcaattgcca
 atattgaacgagcttatcaacgagcatcatcttaacgttaacgtattattttcttctatg
 tcagaaattcaagatgaaacggtttgttattttgtggttgagatttgagcatgatgagtc
 ttaaatgattttaaacttactgattacctttcaaacgacatattcggtatgaggagggtt
 45 atataa

Sequence 3202

VIEFKNVNKFVRKKRETIQALKNVSFKIDQHDIFGVIGYSGAGKSTLVRLVNQLETVSDG
 QVIVDGHEIDTYKEKDLRDIKKDIGMIFQHFNLLNSKSVYKNVAMPLILSKTNKKEIKEK
 50 VDEMLEFVGLADKKDQFPDELSSGQKQRVAIARALVTHPKILLCDEATSALDPATTSSIL
 NLLSNVNRFTGVITIMMITHMSVIQKICHRVAVMENGEVIEMGTVKDVFSHPQNTAKNF
 VSTVINTEPSKELRASFSNRKDSNFTDYKFLDSEQIQLPILNELINEHHLNVNVLFSM
 SEIQDETVCYLWLRFEHDESFNDFKLTDYLSKRHIRYEEVI*

55 Sequence 3203

Contig_0718_pos_1244_585

is similar to (with p-value 6.0e-81)

>sp:sp|P30750|ABC_ECOLI ATP-BINDING PROTEIN ABC. >gp:gp|U70
 214|ECU70214_48 Escherichia coli chromosome minutes 4-6. NID

: g1552727. >gp:gp|AE000129|AE000129_3 Escherichia coli K-12 MG1655 section 19 of 400 of the complete genome. NID: g1786395.

5 atgttttggttcaagtttagattcatctcaattattacaagctctatacgaaacattgtat
atgggtgactgtatcacttgaatcggtgctttaataggtatacctcttggcatcttgta
gtggtaactagaaaaaacggtatatggcgcaatacaatattgcatcaagtattaaatcct
atcattaatatatttaagatcaattccgttcattattttattaatagccatagtgcccttt
actaaattgctagttggcacatctatcggcacacagcagccatagtaccactcacggtt
10 tatgtagcaccttatatcgcacgcttagtagaaaactcattactggaagtcgatgacggg
attattgaggcagctaaagcaatgggtgcatcacctcttcaaattatacgttatttttta
ttgccagaagcacttgggtcattaattcttagctataactacagctattattgggtctcata
ggtagtacagcaatggctggtgctggtggtggcggtataggtgatttggctttagtg
tatggttatcaacgattcgatacaattgtcattgtgattacagtcattgtacttattatt
attgttcaaattatacaaacgcttaggtaactttatcgctagggttatccgtagaaattaa

15

Sequence 3204

MFGSSLDSSQLQALYETLYMVTVSLVIGALIGIPLGILLVVTRKNGIWSNTILHQVLNP
IINILRSIPFIILLIAIVPFTKLLVGTSIGTTAAIVPLTVYVAPYIARLVENSLLLEVDDG
20 IIEAAKAMGASPLQIIRYFLLPEALGSLILAITTAIIGLIGSTAMAGAVGGGGIGDLALV
YGYQRFDTIVIVITVIVLIIIVQIIQTLGNFIARVIRRN*

Sequence 3205

Contig_0719_pos_4886_3567

25 >sp:sp|P52673|CYSI_THIRO SULFITE REDUCTASE (NADPH) HEMOPROT
EIN BETA-COMPONENT (EC 1.8.1.2) (SIR-HP). >gp:gp|Z23169|TRCY
SCOMA_3 T.roseopersicina cysJ, cysI, cysH genes, complete CD
S, and cysB gene 5' end. NID: g1518424.

30 atgaaaaatattaatcatgcagttacttgattctattgctgcatgtggagatgttaatcgt
aatacgatgtgcaatcctaactccttatcaatctcaagtacataaggagattaatgattat
gcaacgcgtataagtaatcacttacttccaagaacaaatgcataatcatgaaatttggctt
gatggtgaaaagggttttagattcgagtgaggaaaaggaacctatttatgggaatacgtat
ttaccacgtaaaattcaaaatagggtattgcagtagccaccatctaataatgatattgacgtctat
tctcaagatatgggtttaatcgctatcggtgaacaagatgagtttaattggatttaattgtg
actatcggtggcggtatgggtatgactcatggtaataactgaaacataatcctcaacttggg
35 cgtctcatagggttttatacctaaggaaaagggtttagatgtatgtgagaaaaatacttaca
atacaacgtgattatggtaatcggtgaaaaatcgaaaaatgcacgtttttaatatagctg
gaccgtctaggagaaacttgggtgactgaagaattaaaccgacgattagggttgggaatt
aaagcgccacgtgattttcgaatttgaacataatgggtgatcgattagggttggattgaaggt
attaataattggaatttcactttatttatacaaaatgggcgtgtgaaagatactgaagac
40 tatttgttaaaaacaaccttaagagaaatcgagaaatccatactggagatttcagatta
tcacctaatacagaacttagttattgcaaatgtttctcctgagaaaaaggaagaatacaa
gctattattgataaatataaaattaacagatggcaaaaattatacaggacttagaagaat
tctatggcttgtgttgctttcccaacgtgtggtttagctatggcagaatctgaaagatat
cttccttcaataattacaaaaattgaagatttatttagatgagtcgtttaaagaggaa
45 gaaataacgattcgtagacaggttgtcccaatggatgtgagagaccagcgctagcagaa
atagcctttatcggtaaagcacctggtaaatataatgtacttaggtggtagttttaa
ggcgaacgtctaaataaaaatatataagagaatatcgacgaaaatgagatatagaaagt
ctacgtccattgttgttgcgttatagtaaagagcgtcttgacggagaaacactttggggac
tttgtaattcgtagcgggtgtgatagccaaagttcatgatggctcgcatatttcatagttaa

50

Sequence 3206

55 MKNINHAVLDSIAACGDVNRNMTMCPNPYQSQVHKEINDYATRISNHLPRTNAYHEIWL
DGEKVLDSSEEKEPIYGNTRYLPRKFKIGIAVPPSNDIDVYSQDIGLIAIVEQDELIGFNV
TIGGGMGMTHTNTETYPQLGRLIGFIPKEKVVDVCEKILTIQRDYGNNRKNARFKYTV
DRLGETWVTEELNRRLGWEIKAPRDFEFEHNGDRLGWIEGINNWNFTLFIQNGRVKDTED
YLLKTTLREIAEIHGTDFRLSPNQNLVIANVSPEKKEEIQAIIDKYKLTGKNTGLRRN
SMACVAFPTCGLAMAESERYLPSLITKIEDLLDESGLKEEEITIRMTGCPNGCARPALAE
IAFIGKAPGKYNMYLGGSFKGERLNNKIYKENIDENEILESLRPLLLRYSKERLDGEHFGD
FVIRDGVIAKVHDGRDFHS*

Sequence 3207

Contig_0728_pos_9983_10327

5 >gp:gp|AF043386|AF043386_2 Clostridium acetobutylicum glyce
raldehyde-3-phosphate dehydrogenase (gap), phosphoglycerate
kinase (pgk), and triosephosphate isomerase (tpi) genes, com
plete cds; and 2,3-bpg-independent phosphoglycerate mutase (p
pgm-i) gene, partial cds. NID: g2829136.
10 atgcagattgaagaaaccttccgagacttgaaaagtcctgcctacggactaggcctacgc
catagccgaacgagcagctcagagcgttttgatcatgctgctaatacgccctgatgctt
caactaacatgttggcttgctggcggttcagctcagaaacaagggtgggacaagcacttc
caggctaacacagctcagaaatcgaaacgtactctcaacagttcgcttaggcattggaagtt
ttgctggcattctggctacacaataacaagggaagactcactcgtggctgcaaccctgctt
actcaaatctattcacacatggttacggttttggggaaattatga

Sequence 3208

MQIEETFRDLKSPAYGLGLRHSRTSSSERFDIMLLIALMLQLTCWLAGVHAQKQGWDKHF
QANTVRNRNVLSTVRLGMEVLRHSGYTITREDSLVAAATLLTQNLFTHGTVLGLK*

Sequence 3209

Contig_0728_pos_8652_7642

is similar to (with p-value 4.0e-59)

25 >pir:pir|I67760|I67760 transposase (insertion sequence IS10
) - Escherichia coli >gp:gp|S67119|S67119_2 BST=somatotropin
...BST/beta-Gal fusion protein [Escherichia coli, LBB84, pla
smid pXT107, IS10L/R-1, PlasmidTransposonInsertionMutant, 3
genes, 1679 nt]. NID: g455674.
30 atggcaattaaagtagcaattaatgggttttggtagaattggctgcttagcattcagaaga
attcaagatgtagaaggtcctgaagtagttgcagttaacgacttaacagatgacgatatg
ttagctcatttattaaaatacagatactatgcaaggtcgcttcactggagaagttgaagtt
atcgaaggtggattccgtgtgaacggtaaaagaaattaaatcattcgatgaaccagatgct
ggtaaattaccatggggcgatttagatatcgacgtagtattagaatgtactgggttctat
actgataaagaaaaagcacaagctcacatcgatgcaggtgctaaaaaagtattaatctca
gctccagctaaagggtgatgtaaaaacaatcgattcaacactaacatgatacattagat
35 gggttcagaaacagttgtttcaggtgcttctgtactactaactcattagcaccagttgca
aaagtttttaagtgaagaaattcggttttagttgaaggtttcatgactacaattcacgcttac
actggtgacaaaatacacagaacgcacctcacagaaaagggtgacaaacgtcggtgcacgt
gcagcagctgaaaatattatccctaactcaacaggtgctgctaaagctatcggttaaagtt
attccagaaatcgatggtaaattagacggtggagcacaacgtgttcagttgctactggt
40 tctttaactgaattaactgtagtatttagacaaacaagatgtaactgttgaccaagttaac
agtgcattgaaacaagcttctgacgaatcattcggttacactgaagacgaaatcgatct
tctgatattgttggtatgacttacggttcattatttgatgcgactcaaactcgtgttatg
actgttgagatcgctcaattagttaaagttgcagcttggtacgacaatgaaatgtcttac
actgctcaattagtagctacattagctcacttagctgaacttttctaaataa

Sequence 3210

50 MAIKVAINGFGRIGRLAFRRIQDVEGLEVVAVNDLTDDMLAHLKDYDTMQGRFTGEVEV
IEGGFRVNGKEIKSFDEPDAGKLPWGLDIDVVELECTGFYTDKEKAQAHIDAGAKKVLIS
APAKGDVKTIVFNTNHDITLDGSETVVSASCTTNSLAPVAKVLSDEFGLVEGFMTHAY
TGDQNTQDAPHRKGDKRRARAAENIIPNSTGAAKAIGKVIPEIDGKLDGGAQRPVATG
SLTELTVVLDKQDVTVDQVNSAMKQASDESFGYTEDEIVSSDIVGMTYGSFLDATQTRVM
TVGDRQLVKVAAWYDNEMSYTAQLVRTLAHLAELSK*

Sequence 3211

Contig_0732_pos_1186_2424

55 >gp:gp|AF043609|AF043609_1 Arthrobacter viscosus aluminum r
esistance protein (Alu-2) gene, complete cds. NID: g2827438.

atgcaagatttttagcaatttagttgaagaagttgaaaacacacttattccttacttttaga

aaaattgaaaagcgtgcattatttaatacaggaaaaggtcttaaatagcttttcaccatggt
 aaagctagcgaagtgatttacaggggtctacgggttatggatatgatgattttgggaga
 gaccatttagaacaatttatgcgcacacatttaaagcagatgacgcacttgtaagacct
 caaattatttcagggtactcatgctatttacttttagctttacaaagtacgttaaaaaacaat
 5 gatgaactactttatattacaggtagtcacatgatacacttctagaagtcattggtata
 aatggcaatgggtgttgaaagcttaaagaatatgggtgttcgctataatgaagtcgaatta
 cgtgacgggtcgaattgatattcctaaagtcactgcaattaatgacaatacaaaaagtt
 gtagcaattcaacgatcaaaaaggatgatcaacgtccatcaattacaattaatgaaatt
 gaacaagcaataacatctattaaagaggtttatcccaatatcattattttgttgataat
 10 tgttatggagaattttagagaagataaagaaccgattgaagtaggtgctgatttaacgcgc
 ggatcatttaataaaaaatccaggtggaggttttagctaaaaattggaggatatattgctggt
 agacaagacttaattgaacgctgtggttatcgtttaacagcaccaggcatggtaaggaa
 gcaggagcctcacttaattctttacaagaaatgtatcaaggattctttctagcgccacat
 gtggttagccaaagtttaaaaggtgcactgtttactagttgttattagaaaaataaac
 15 atgaagacctccccataatataatgtttatcgtagacacttaattcaaacgggttcaattt
 gagaccaaagagcaaatgatttcattttgccaaagtatacaacacgcttcaccaattaac
 gcacatttttagtcagaaacctagctatatgcctggatacgaagatgatgtcatcatggct
 gcaggtagcttttatcagggctcgctcattgaattatccgcagacggacctatacgtccg
 ccttatgaagcatatgttcaaggtggtttaacttatgaacatgtcaaattagctgttaca
 20 cgtgcgggtgcaacatatgcaagaaaacaatttactataa

Sequence 3212

MQDFSNLVEEVENTLIPYFRKIEKRALFNQEKVLNAFHVKASESDLQGSGTGYGYDDFGR
 DHLEQIYAHTFKADDALVRPQII SGTHAITLALQSTLKNNDellyITGSPYDTLLEVIGI
 25 NGNGVESLKEYGVRYNEVELRDGRIDIPKVITAINDN TKVVAIQRSKYDQRP SITINEI
 EQAITSIKEVYPNIIIFVDN CYGEFVEDKEPIEVGADLIAGSLIKNPGGLAKIGGYIAG
 RQDLIERCGYRLTAPGIGKEAGASLNSLQEMYQGFFLAPHVVSQSLKGALFTSLLLEKIN
 MKTSPKYNVYRTDLIQTVQFETKEQMISFCQSIQHASPINAHFSPEPSYMPGYEDDVIMA
 AGTFIQGSSIELSADGPIRPPYEAYVQGGTYEHVKLAVTRAVQHMQENLL*

Sequence 3213

Contig_0732_pos_2905_3270

is similar to (with p-value 5.0e-55)

>gp:gp|X76490|SAGLNAR_2 S.aureus (bb270) glnA and glnR gene
 35 s. NID: g1134885.
 atgtctaataatgattcaatcagacgaacatggccggttttctctatgagtggttagtaaa
 ttgacagatttatcaccaagacaaatcggttactatgaaacacatgaacttgatgac
 gaaagaacagatggaaataagagattattttctatgaacgatttagagaggttgtagaa
 ataaagtcctcttatcgaaaagggtttaataattagaggtattaaacaaatttatattcgat
 40 gagcaagggtcatttaactactgatgaacaagagacaagaaagagaatgattgttgacgca
 acgcagaaaccacgtagtgaaacattaccaataaatcgcggtgatttatctcgatttatt
 aatga

Sequence 3214

45 MSNDSIRRNMAVFSMSVSKLTDLSPRQIRYYETHELVMPERTDGNKRLFSMNDLERLLE
 IKSLIEKGFNIRGIKQIIFDEQGHLLTDEQETRKRMIVDATQKPRSETLPINRGDLRFI
 K*

Sequence 3215

Contig_0733_pos_1000_1344

is similar to (with p-value 2.0e-46)

>sp:sp|P94453|ALF_BACST FRUCTOSE-BISPHOSPHATE ALDOLASE (EC
 4.1.2.13) (FRAGMENT).
 50 atgaaagaatgttaatcgatgcgaagaaaacgggttatgcggttggtcaatacaatctt
 aataacctcgaatttacacaagctattttagaagcgtctcaagaagagaatgcgccagtt
 attttaggtgtttctgaaggggcagctcggttatatgagtggtttttatcacagttgtgaaa
 atggtagaaggtttaatgcatgacttaacatcacatcccagtagcaattcatttagac
 caggttcaagctttgaaaaatgtaagaagcaattgatgctggattcacatctgtaatg
 55 attgatgcatctcatagtccttttgaagaaaatgttgaaatatag

Sequence 3216

MKEMLIDAKENGYAVGQYNLNNLEFTQAILEASQEENAPVILGVSEGAARYMSGFYTVVK
MVEGLMHDNLNITIPVAIHLDHGSSFEKCKEADAGFTSVMIDASHSPFEENVEI*

5

Sequence 3217

Contig_0733_pos_12244_11498

is similar to (with p-value 4.0e-40)

>sp:sp|P55476|NODI_RHISN NODULATION ATP-BINDING PROTEIN I.
10 >gp:gp|AE000076|AE000076_5 Rhizobium sp. NGR234 plasmid pNGR
234a, section 13 of 46 of the complete plasmid sequence. NID
: g2182419.

gtgagtggtcactcattttatcggtgtgaggaggaacaatttatgatagaggtaaaaaat
gtaagtaaataccttttggtaaacaacaagtggtagatgatatactatattcatttaagtgc
15 ggtgaggttgtgggtctaatacgtccttcaggtactggaaaaacgacattaatacagtcg
atatttaggcattggagaaaaattgatgggtgggcaagtcactattcaagaacatacaatgccg
aatagaaaagattatcaataattgggttatatggctcaaaatgatgctttatataatgat
ttaactggacgtgaaaatttaacgttttttcgcaagaatttatatgcgtgataaagaagat
attaaaaaacgtgtgaacctatgcagttccatgggtcaattagacaatgatttagataag
20 aaagttgaaatgtattctggtggaatgaaacgacgcttatcttttagctattagcttttta
caaaatcctaatactccttatattagacgaacctacagttggcattgatcctaaattgcgt
cagacgatttgggaaggatttaactaaagcaaaggctgaagataaatgcattatagtgaca
acgcatgtattagatgaagctacacgctgtgataagctcgtattaatgaatcaaggaaag
atattggcaacgggtacaccagatgaagtgaaaaaacaatatcacacagatacgattgaa
25 ggcgtatttctgaatatggagggataa

Sequence 3218

VSGHSFYRCBEEQFMIEVKNVSKSFGKQQVVDDISISFNAGEVVGLIGPSGTGKTTLIQC
ILGMEKIDGGQVTIQEHTMPNKRKILSNIGYMAQNDAALYNDLTGRENLTFFARIYMRDKED
30 IKKRVNLCSSMVQLDNDLDKKVEMYSGGMKRRLSLAISFLQNPINILDEPTVIGIDPKLR
QTIWKDLTKAKAEDKCIIVTTHVLDEATRCDKLVLMNQGKILATGTPDEVKKQYHTDTIE
GVFLNMEG*

Sequence 3219

35 Contig_0733_pos_9564_7870

is similar to (with p-value 7.0e-31)

>gp:gp|AB009866|AB009866_7 Bacteriophage phi PVL proviral D
NA, complete sequence. NID: g3341907.

atggcaagattagggattatgttacagaatatgcaaaaaagtagttaatggcgatatt
40 atagctagtaaaaaaacgtgaaagcctgtcaacgccatttagatgacttgaacgattcg
gaactcccttatcattttgatgtaaagaaagctaatacattattaagtttcttgaaatg
ttgccagatcctaaaactggttaacaattatcggttaggcggttttcaaaaattcattgct
ggtagcttaaatggttggtacgacagacatgggtacaaaagatttacaaaagcctatata
tcaatgagcagaaaaaatggttaaacattattgatctctggaatggcattgtacgattta
45 ttgatgggtaaagatccgttgaaatgaacgggttgattgggttgagcgccaattcaagagac
caagctggtatagcatagcatatgacattggcacaactgaaagctattagaagcgtttct
cctaagggttaaatcgatgactaagataacgccaagtgcataaagaaatattgaatattaat
gatcgaagtaaagttaaagccgtttcaaatgaagctgcaaattagaaggctcatcagttt
agctacgcaatcatcgatgaatatcatgaagctaagataaaaagatttatgaaacgtta
50 agacgtgggcaagtgctactgcacaacctatattaattattatctcaacagctggaact
aatttgaatgggtccgatgtatgaagaattttatatttgataagatacttgacggcata
gcaaaaaatgaaaactactttgtttctgtgctgaacaagatgatgagaaagaagtatat
gacgttaaaaacttggattaaatccaatccacttatggagttgccagaaatggcacaattg
ttaactaagaatattcaaccagaagttaaaactgcaattgatagtggttcaggattaaat
55 gggatattaataaagaatttcaatatgtggcgtgcagcaagcacagaatcttatttagat
ttcaatgattggaagaaaaatgaaatagactttgatataaatggctctaaaacttatatc
ggtttagacttatcgctgctgacgacttaaccgcagtatcgtttgttcattcttgatgaa
gataatcaagagtattatgtaactagtcattcgtttgtggctactaaagggtggattagat
ggcaagattgatagagactttattgattacagacaacttgcaaaaagtggttattgtacg

attaccgatttacaaagtgaattatcaatactgaccaagttttaaattacattgagaat
tatatcgaccaatataaattagacgtacaagcgttatgttatgatccttactcaatacat
gggtgttattgcagaaattgagcgtagagattggccttatgatttagtagaaatcagacaa
gggccacaaacactatctaattccgatactggatttttagactgaaagtgattaatggggac
5 atcaagcatcataaaaaatccgcttactagacattgcagtcaaaaatgctgtggcaaaagat
accaatgactcattaatgattgaaaagaagatgaaccgagaaaaaatagatccactcatg
gctaccataatttgccttatgttatggccttgtgaacatgaatgggacacagaaactttaatg
ccattgttcttatag

10 Sequence 3220
MAKIRDYVTEYAKKVVNGDI IASKKNVKACQRHLDLNDSELPYHFDVKKANHI IKFLEM
LPDPKTKGKQLSLGGFQKFIAGSLNGWYDRHGYKRFKAYISMSRKNKGTLLISGMALYDL
LMGKDPLNERLIGLSANSRDQAGIAYDMTLAQLKAIRSVSPKVKSMTKITPSAKEILNIN
DRSKVKAVSNEANLEGHQFSYAIIDEYHEAKDKKIYETLRRGQVLLHNPILIIISTAGT
15 NLNGPMYEEYLYIDKILDGIAKNENYFVFCAEQDDEKEVYDVKTWIKSNPLMELPEMAQL
LTKNIQPEVKTATIDSGSLNGILIKNFMWRAASTESYLDNFNDWKKNEIDFDINGSKTYI
GLDLRADDLTAVSFVHLDEDNQEYVYTSHSFVATKGGLDGKIDRDFIDYRQLAESGYCT
ITDLQSGIINTDQVLNYIENYIDQYKLDVQALCYDPYSIHGVIAEIERRDWPYDLVEIRQ
GPQTLSNPILDRLKVIINGDIKHHKNPLLDIAVKNNAKDTNDSLMIKKMNREKIDPLM
20 ATIFAYVMACEHEWDTETLMPLFL*

Sequence 3221
Contig_0733_pos_7622_6432
is similar to (with p-value 2.0e-16)
25 >gp:gp|AB009866|AB009866_5 Bacteriophage phi PVL proviral D
NA, complete sequence. NID: g3341907.
atgaacagagatttagaacgattattgtattggcaagaacatggcacacatgcaagctat
gttggtataaacgcgctacgtaacagtgatgtatttactgctacacgtattatatctgca
gacattgcaagtaccaagttgaaagttaaaggtcacgaaacaaatcacgtgatggaccaa
30 atactggatctatttaataacaatccgtattcggacttaccgggttggcactttaagttt
ataatcatcgcaaatatgctgcttaacgggtcaatcttttgttgaaattgtgcgtggcaaa
aatgattttcctgttggattccacttcttacataacgacttagtaggaattgaggaaaaa
gacggcgaaattatttacaacgtaagtgaagatgtggaaggtaatgccgttaagataaca
agcgatgatataattacatttcagatatatcacattagatggatatataggatacagtcg
35 ttgtatgcactagcacatgagatttggatttctcaaggctcctaagagcttccctgcgtaac
ttcttcgataatgggtgggacttcgacatcagatttgaagtatagaaaagggcaaatcaat
gctgaacaattaagagaattgaaaaagaacttttcagaaagtcattaaaaaacaacggt
ggttttagttgctatcgatgacacaatggaatttaacagactacaaattcctaccgaagta
ttgaacttcttaaatagttataagttcagcacatctcaagttgctaaagcggttcggttg
40 ccggtatctaaactaggtattgaaacagtcataacatctatcacacaagcaacttagag
tatttgcgaagtacattagatccaatatttaaaatgatgattgctgaactcgaaacgaaa
atatttaaatttattgattctggtaacgaattagagtttgactcatcacgtctcattgac
attgatccagagttacaattacaacgtattactgaattgcatagtaaaggaattatttca
acagacgaagctagaagtgtatttggctatcaacctattgaacatggcgagcaaccattg
45 gttgatcttaacagagcgccacttaacactttagaaaattaccaaaaaatcgaaaattgac
aaagaagtcgaaaagaactccattaaagggggtgatgagtatgacgaatag

Sequence 3222
MNRDLERLLYQEHGTHASYVGINALRNSDVFTATRIISADIATKLVKGHETNTVMDQ
50 ILDLFNNNPYSDLPGWHFKFII IANMLLNGQSFVEIVRGKNDFPVGPHFLHNDLVGIEEK
DGEIIYNVSEDEVEGNAVKITSDDILHFRYITLDGYIGYSPLYALAEIGISQGSKSLRN
FFDNGGTSTSVLKYRKQGINAEQLRELKKNFSESQLKNNGLVAIDDTMEFNRLQIPTEV
LNFNLNSYKFSTSQVAKAFGLPVSKLGIETVNTSITQANLEYLQSTLDP IFKMMIAELET
IFKFIDSGNELEFDSSRLIDIDPELQLQRITELHSGKIIISTDEARSVFGYQPIEHGEQPL
55 VDLNRAPLNTLENYQSKIDKEVEKNSIKGGDEYDE*

Sequence 3223
Contig_0733_pos_6406_5885
is similar to (with p-value 2.0e-46)

>gp:gp|AB009866|AB009866_4 Bacteriophage phi PVL proviral DNA, complete sequence. NID: g3341907.

atgggtgtcgaagggttacgccattatctttaattcaatgagtgatgatttgggtggattt
 5 agagaaattgtagcgcctaattgctttaaatgatgtagatgtaagtgatgtgaaatgtcta
 atcaatcatgatttttagttatgttataggacgcacacaagcaggaacgcttgagctacag
 gtggatgaaaaagggtatatactttaaatgccacttacctaatacatcatacgcaagagat
 atttatgagaatattaaagcaggcaacggttaatcagtcagtttcttttacacattgccca
 cctaatactcaacggctcgtacgtggcaaaacatagataatgagtcagttcaaacata
 aataaaatcgatgaattgattgaggttagtattgttacagtgccagcctacaaagataca
 10 tcggttgaagtcggtcaacgtgcgaaagacttaaaagaattcaaacagttggaacaaatg
 aagatagcattggatttagaaagcctacgttttgaaacgtaa

Sequence 3224

MVVEGYAIIFNSMSDDLGGFREIVAPNALNDVDVSDVKCLINHDFS YVIGRTQAGTLELQ
 15 VDEKGLYFKCHLPNTSYARDIYENIKAGNVNQC SFFYTLPPNDSTARTWQNIIDNEYVQTI
 NKIDELIEVSIVTVPAYKDT SVEVGQRAKDLKKFKQLEQMKIALDLES LRFET*

Sequence 3225

Contig_0733_pos_5844_4492

>gp:gp|X97563|BPHA3GP3_5 Bacteriophage A2 gp3 gene and 4 op
 20 en reading frames. NID: gl523807.
 atggctaatttagatgagcgcacaaaagaaatcgctaattctgattttctaaagcgcaagaa
 gcagtcgaaaaaaggcgacctcgaaactgctcgtaattttaaagctgatatgatgctcaa
 aagaaagagtacgaagaactcgaaacagctttcaaaagaaattgaagcgtcagcacctaaa
 25 caagatgaaccacctaagatgaaggtgcagaagttgaagataacaaagatggtaattct
 ggagaagaatcagagaacaaaccttctgatgatgaaccagaaggaacttcagatgaagaa
 aaacctgatgatgcacaaaaccagatgacaaacctgaagaaacaccagaaacacctact
 attgaaaaagtagaagaaccaacagaagaagaattaaaaaaagaaaaagacaaaaaagaa
 ggagcgaacagttcttatggctaaattaaaccaaactccagagacaaacgaagaattcta
 30 gcatcttgaaacagtcacatgaaatcaaaaggggctaaacgtgacaatgttaaatctgatgac
 gttggcgtaactatcccagaggatattaaatatattcctgaaaaagaagttaagacagtc
 caagacttatcagaatttggtacaaaaacttcagtatcaactgcaagtgggaaataccg
 atcttaaacgtgctaacgctaaattcaacactgttgctgaattagagaaaaacacctgag
 ttagctcgtccggaattcgaaacaatcaattgggaagtagacacttatcgtggatctatt
 35 ccgatttcacaagaagcatttagatgattcagttgctaacttaactgctattgtttctgaa
 aatattaacgaacaaaaaatcaacactttaaatgaacgtatttggtgaagttttaaagca
 ttcaatcctactagtgtttctaattgttgacgacttaaaagaaattatcaacgttaatta
 gatcctgggttatgaccgccaattatctgtactcaaagtttctatcaaaaactagataca
 ttaaagatggtaacgggtcggtatttactacaagacagtatcatcaacactgcaggtaac
 40 actgtgttaggtatgaatgtaacagttgtgctgatgacttggttaggtaaaaatggagat
 gcattagcatttatttggtgatgtaaacgcgggtgtgttatttgcagaccgtacagacgtt
 tctgttcaatggattgaaaatgaaatctacggtaaatacttaattgggtgctttccgtttc
 gatgtgaaacaggctgataaaaaatgctggtttcttcgtaacatttgaagagcgtttatat
 tacttcatattgggcaatggatgtatacagatga
 45

Sequence 3226

MANLDERKKEIANLISKAQEAVEKGDLETARNLKADIDAQKKEYEELEQLSKEIEASAPK
 QDEPPKDEGAIVEDNKGNSGEESENKPSDDEPEGTSDEEKPDDAPKPDDKPEETPETPT
 IEKVEEPTEEELKKEKDKKEGAKRSMAKLNQNPETNEEILAFEQYMKSKGAKRDNVKSDD
 50 VGVTPEDIKYIPEKEVKTVQDLSELVQKTSVSTASGKYPIILKRANAKFNTVAELEKNPE
 LARPEFETINWEVDTYRGSIPISQEALDDSVANLTAIVSENINEQKINTLNERIGEV LKA
 FNPTSVSNVDDLKEIINVKLDPGYDRQIICTQSFYQKLDTLKDGNGRYLLQDSIINTAGN
 TVLGMNVTVVRDDLKGNGDALAFIGDVKRGVLFADRTDVSQWIENEIYKGYLMGAFRF
 DVKQADKNAGFFVTFEERLYYFILNGCIR*
 55

Sequence 3227

Contig_0734_pos_2701_2249

is similar to (with p-value 9.0e-25)

>sp:sp|P00937|TRPG_YEAST ANTHRANILATE SYNTHASE COMPONENT II

(EC 4.1.3.27) (CONTAINS: GLUTAMINE AMIDOTRANSFERASE; INDOLE
 -3-GLYCEROL PHOSPHATE SYNTHASE (EC 4.1.1.48) (PRAI)). >pir:p
 ir|S38049|NNBY2 anthranilate synthase multifunctional enzyme
 - yeast (*Saccharomyces cerevisiae*) >gp:gp|X75951|SC6ORF_9 S
 5 .*cerevisiae* URA1, SAC1, RSD1 and TRP3 genes and 6 new orfs.
 NID: g473130. >gp:gp|Z28211|SCYKL211C_1 *S.cerevisiae* chromos
 ome XI reading frame ORF YKL211c. NID: g486376.
 gtggtggttaatacatctacattttattgcattatcgctcagtatggacaagctggtgcatct
 attatttttattaatagtaaataattttaagtgtatgaccaattaaaagaattgtattcatat
 10 gcaacaaaccataatttagaagctctagtagaagttcatacaatttagagaacttgaacgt
 gcacaccaaatcaaccctaaaattattggtgttaataatcggtgattttaaaacgatttgaa
 accgatgttctacatacaaaataaattacttaagtttaaaaagtctaattgctgtacatt
 tcagagagtggcattcatacaaaaagaagatgttgagaaaatagtagattcaagattgac
 ggtttactgttaggggaggcattaatgaaaacaaatgacttaagtcagtttttgcctagt
 15 ttaaagttaaagaagaatctctatgatagtaa

Sequence 3228
 VVVIHLHLHYRQYGQAGASIILLIVNILSDDQLKELYSYATNHNLEALVEVHTIRELER
 AHQINPKIIGVNNRDLKRFETDLHTNKKLLKFKKSNCYIYESGIHTKEDVEKIVDSSID
 20 GLLVGEALMKTNDLSQFLPSLKLKKNLYDS*

Sequence 3229
 Contig_0737_pos_1135_1560
 is similar to (with p-value 2.0e-19)
 25 >gp:gp|AF001974|AF001974_1 *Thermoanaerobacter ethanolicus* p
 utative TrkG gene, partial cds, and putative TrkA, xylose is
 omerase (xylA) and xylulose kinase (xylB) genes, complete cd
 s. NID: g2581794.
 gtgagaccgactgtaccgaatgctgatacgacttcgaagagtattttgattaaaggtata
 30 ttggaattgattatggtaagtataaacgtaaccataccgataaatgcaatagaaatgaga
 atagttacgaaagacaactgtatgtatctttctgtatatttctctattaaatatagagttg
 tttttttctttgctgtatcgtattaaaaatagcgattgttgcaataacaaaagttgtaacc
 tttataccacctgcagcactcaatggtgcacctccaataaacatgagagccataagtaaat
 aaagctgtcgggtgttttaattgttccaacgtcaattgtgttaaatcctgcagtccttgtt
 35 gtcactgattggaaaaatgcatttcctattttttcaattaateccatgtgtaacatagag
 ttttga

Sequence 3230
 VRPTVPNADTTSKSILIKGILELIMVSINVTIPINAIEMRIVTKDNCMYLSDISLLNIEL
 40 FFSLRIVLKIAIVAIVTKVVFIPPAALNGAPPINMRAISNKAVGVLMPFITSIVLNPVAVLV
 VTDWKNAPFIFSINPMCNIIEF*

Sequence 3231
 Contig_0737_pos_1585_2199
 45 is similar to (with p-value 3.0e-19)
 >sp:sp|P43440|NTPJ_ENTHR V-TYPE SODIUM ATP SYNTHASE SUBUNIT
 J (EC 3.6.1.34) (NA(+)- TRANSLOCATING ATPASE SUBUNIT J). >g
 p:gp|D17462|ENENTP_11 *Enterococcus hirae* ntp genes for Na+ -
 ATPase subunits, complete cds. NID: g487271.
 50 gtgcctatcaaaaattaataatccagtggttagttaacacgagtttagagtgttaaggataat
 tttctaaaactttttggcggttcacaaatcaaccacgactaaatgtccaaacctcccaaa
 atgataagatttggaaatagtataatgattaccggatcatttgaaaaatcgattaagttg
 tttttaaaaagggcggaatcctgcgttggttaaatgcggaaactgaagtgaataaacttaaa
 aatagacctttacctatgccaaattttggaataaacgataaacatagacaaagtgtacca
 55 aataattcagtgggcgatgctgtatatggctagatgtttaataagtttaattacaccaccg
 ggttcgtcaatattccaagtaatcataaataaaattctattgttaattgatattttctta
 tttaaaaagatgagagtttagcattgctacagtgacaatacctaattccaccaatttgtata
 agtaataagataatgatttctccaaaaatattaaattgtgttccaacatcaactggtgat
 agacctgttactgtgaatgcgctagaagctacaacaaatgcgtcaataaagtttaataggt

ttcttccctgtatag

Sequence 3232

5 VPIKINNPVVNTSLECKDNFLKLLAFHKSTTTKCPKPPKMISIGIVIMITGSFEKSIKL
FLKRANPALLNAETEVENKLNRLPMPNFGINDKHRQSVPNNSVAMLYMARCLISLITPP
GSSIFQVIINKILLIDIFLFFKKMRVSIATVTIPNPPICISNKIMISPKILNCVPTSTGD
RPVTVNALEATNNASIKLIGFFPV*

Sequence 3233

10 Contig_0737_pos_2249_978

is similar to (with p-value 1.0e-54)

>gp:gp|D89592|D89592_3 *Vibrio alginolyticus* rhIE, KtrA and
KtrB genes, complete cds. NID: g3927863.

15 atgttatttctgttgacaactttaattggtgcttttctactctatttgcctatacaggg
aagaaacctattaactttattgacgcattggtttagcttctagcgcattcacagtaaca
ggtctatcaccagttgatgttggaacacaatttaatttttggagaaatcattatctta
ttacttatacaaattggtggattaggtattgtcactgtagcaatgctaactctcatcttt
ttaaataagaaaatatcaattaacaatagaattttatttatgattacttggaaattgac
gaaccgcggtggtgtaattaaacttattaaacatctagccatatacagcatcgccactgaa
20 ttatttgggtacactttgtctatgtttatcggtttattccaaaatttggcataggtaaagg
ctatttttaagtttattcacttcagtttccgcatttaacaacgcaggattcgccttttt
aaaaacaacttaatcgattttcaaatgatccggtaatcattatcactattccaatactt
atcattttgggaggtttgggacatttagtcgtggtgatttggggaacgcaaaagt
agaaaattatccttacactctaaactcgtggttaactaccactggattattaattttgata
25 ggcacggttttcttcttttactagaaaatcaaaactctatgttacacatgggattaatt
gaaaaaataggaaatgcatttttccaatcagtgacaacaaggactgcaggatttaacaca
attgacgttggaacattaaaacaccgacagctttattacttatggctctcatgtttatt
ggaggtgcaccattgagtgctgcaggtggtataaaggttacaacttttggtattgcaaca
atcgctatttttaatacagatacgcaagaaaaaacaactctataatttaatagagaaata
30 tcagaaagatacatagcagttgtcttctgtaactattctcatttctattgcatttatcggt
atggttacgtttataacttaccataatcaattccaatatacctttaatcaaaatactcttc
gaagtcgtatcagcattcggtacagtcggtctcactatggatttaacttccgaatactat
aattggactgagtttattattatcatcgtaaatgttatgtggtaaaattggattactgaat
attagtagagcgcttgttccacctaaagaccctaaaaattatagatataccaaggacac
35 attcacttataa

Sequence 3234

40 MLFLTLTLIGAFLLYLPYTGKKPINFIDALFVASSAFTVTGLSPVDVGTQFNIFGEIIL
LLIQIGGLGIVTVAMTLFLFLNKKISINNRIIFMITWNIDEPGGVIKLIKHLAIYSIATE
LFGTLCLCLSFIPKFGIGKGLFLSLFTSVSAFNNAGFALFKNNLIDFSNDPVIIITIPIL
IILGGLGLHVVDLWNAKSFRLKSLHSLKLVLT'TTGLLILIGTVFFPLENQNMLHMLGLI
EKIGNAFFQSVTTRTAGFNTIDVGNIKTPTALLLMALMFIGGAPLSAAGGIKVTTFVIAT
IAIFNTIRKEKNNSIFNREISERYIQLSFVTILISIAFIGMVTFILTIINSNIPLIKILF
EVVSAFGTVGLTMDLTSEYYNWTEFIIIVMLCGKIGLLNISRALVPPKDPKNYRYTKGH
45 IHL*

Sequence 3235

Contig_0738_pos_1259_24

50 >sp:sp|O34863|UVRA_BACSU EXCINUCLEASE ABC SUBUNIT A. >gp:gp
|Z99122|BSUB0019_13 *Bacillus subtilis* complete genome (secti
on 19 of 21): from 3597091 to 3809700. NID: g2636029. >gp:gp
|AF017113|AF017113_13 *Bacillus subtilis* 300-304 degree genom
ic sequence. NID: g2618830.
55 atgcgtgatttaggtaatacacttattgtcgttgaacatgacgatgatactatgagagca
gctgattatttagttgatgtgggtccgggagctggtaaccacggtggagaggttgtctca
agtgttacccttaataaagtaataaagataaaaaatccttaactggcattatattaagt
ggaaaaaacgaattgaagtccctgaatacagacgagaaatcacccgatagaaagattcaa
attaaaggtgctaaaagtaataatttgaaaaatgtaaatgtagacttccactatctgtc
ttaactgtgttacaggtgtgtcaggttctggtaaaagttcactcgtcaatgaaatttta

tataaagcattagctcaaaaaattaataaatctaaagtgaacacctgggaattttgatgaa
 attaaaggaattgatcaattagataaaaatcattgatattgatcaatcgccaataggtaga
 acaccacgttcttaacccagccacatacactgggtgtctttgatgacataagagatgtcttt
 gcacaaacgaatgaagctaaaaatacagggttatcaaaaaggtagatttagttttaatgtc
 5 aaagggtggacgatgtgaagcttgtaaaggatggaattataaaaaattgaaatgcatttt
 ttaccagatgtctatgtaccttgtgaagtatgtgatggtaaacgctataatcgtgagact
 ttagaggtaacatacaaaggtaaaaaatattgcggatgtattagaaatgactgttgaagaa
 gctacgcatttctttgaaaatattcctaagattaaacgtaattacaaacacttgttagat
 gttgggttgggtacattacttttaggtcaacaagctactacattatctggtggcgaagcg
 10 caacgtgtaaaaactcgcacagaattgcacaaacgttcaacggggcggttctatttatatt
 cttgatgaaccaactacaggattacatgtcgacgatataagtcggtttatataagggtattg
 aatcgtatagtggaaaatgggtgatacggctcgttattatcgaacacaatcttgatgttatt
 aaaacggctgatcatattattgatttaggtccagaaggcgggtgaagggtggaggaacaatc
 atcgcaactggtacacctgaagagattgctcaaaaataaagggtcttacactgggtcaatac
 15 ttaaaaccagtatttagagagagacagcgttgaatag

Sequence 3236
 MRDLGNTLIVVEHDDDTMRAADYLDVVGPGAGNHGGEVVSSTPNKVMKDKKSLTGQYLS
 GKKRIEVPYRREITDRKIQIKGAKSNLKNVNVDFPLSVLTVVTGVSGSGKSSLVNEIL
 20 YKALAQKINKSVKPGNFDEIKGIDQLDKIIDIDQSPIGRTPRSNPATYTGVFDDIRDFV
 AQTNEAKIRGYQKGRFSFNVKGGRCCEACKGDGIIKIEMHFLPDVYVPCVECDGKRYNRET
 LEVTKGKNIADVLEMTVEEATHFFENIPKIKRKLQTLVDVGLGYITLQQAATTLSGGEA
 QRVKLASELHKRSTGRSIYILDEPTTGLHVDDISRLKVLNRIVENGDVTVVII EHNLDVI
 KTADHIIIDLPEGGEGGGTIIATGTPEEIAQNKGSYTGQYLKPVLERDSVE*

Sequence 3237
 Contig_0741_pos_5345_6325
 is similar to (with p-value 1.0e-21)
 >sp:sp|P14940|ADH_ALCEU ALCOHOL DEHYDROGENASE (EC 1.1.1.1).
 30 >pir:pir|A30196|A30196 alcohol dehydrogenase (EC 1.1.1.1) -
 Alcaligenes eutrophus >gp:gp|J03362|AFAADH_1 A.eutrophus al
 cohol dehydrogenase (ADH) gene, complete cds. NID: g141899.
 atgtttaaaaagattgctactataatagggttcgacattatttggtacagttttatttcgca
 aaagtgaagaaagcgtagttataaaagttttttacaagagaaaatgattagaatatca
 35 ggaatgaaaaagacatttgaaagtatagatgacgcgaaaaaagctttgaatgagactaaa
 tatcaaaactgcaggtaaatataatggaacaacatatgaatttaagcataaagttcaata
 agagattattatgggtctctttagctctatgttgtaaatgatcattggtcttccagatcaacgc
 acggtcttatatgtacatggaggcgcatgggttccaagatcctttggaaaatcattttgaa
 tatttagacttactcgttgatgcgctcgatgctagggtgattatgcccgatatatcctaaa
 40 ataccacacagagattatcgtagacatttgaaattatacaaaaaatatataagcgatta
 ttgactaaaattgatgaacctgaaaacttgggtcatcattggagattcagccgggggacaa
 attgcattagcttttgcacaaatgttaaaaaaagagcaactcagtcacactggccatatt
 gttcttatttaccgggtgcttgatgcgacatttaagaatccagaagcaagaaaatatgaa
 aaagaagatccaatgcttggaaattgaaggcagtaaatatctttagtaggttatgggctggt
 45 gatgcaccactagatgactataagatgtctccaatgaatgggtgatttagaaggcctagga
 catattacacttactgttaggaaccaaaagaacattatatcctgatgcagtttaagttctct
 cacatgtttaatgataaaggaataaagcatcagtttatcccagggttacaattttatttcat
 atttatcccttattccctatcccagagcgtcaacgcttttttagaacagcttaaaaaaatc
 attgtcacaaaagagttataa

Sequence 3238
 MFKKIATIIGSTLFGTVLFAKVKEKRSYKSLQEKMIRISGMKKTFFESIDDAKKALNETK
 YQTAGKYNGTTYEFKHKVQIRDYVYVNDHGLPDQRTVLYVHGGAWFQDPLENHFE
 YLDLLVDALDARVIMPVYPKIPHRDYRTTFELLTKIYKRLLTKIDEPENLVIIGDSAGGQ
 55 IALAFQMLKKEQLSQPGHIVLISPVLDAFKNPEARKYEKEDPMLGIEGSKYLVELWAG
 DAPLDDYKMSPMNGDLEGLGHITLTVGTKETLYPDAVKFESHMLNDKGIKHQFIPGYNLFH
 IYPLFPIPERQRFLEQLKKIIVTKEL*

Sequence 3239

Contig_0741_pos_3244_2630

is similar to (with p-value 2.0e-21)

>gp:gp|L38252|ACCEST_2 Acinetobacter lwoffii orf1 and ester
ase (est) genes, complete cds. NID: g1209221.

5 atgttatcagatatacttccaacaggttatgaaattgggtgttttaaaaggtaaagttaaa
cctggctgtacagtagccattgttaggtgctggctcctgtagggttagcagcattacttaca
gcacaattctattcaccttcaaaaattattatgattgatttagatgataatagattagaa
accgctaagaactaggtgtacgcatttaattaactctaaagagactgaaaccgcaatt
aaaaaggtaaaatcgtaaataccacgtgggtgttgatgttgctattgaagctgtcggaatt
10 ccacaaacctttgatttatgtcaaaaatttaattgggtgtcgatgggtacgattgctaattgtt
gggtgtgcattgggttacctgtacaacttgatagataaattatggattaaaaatattaac
gtaactactcgtgttagtttcaggaaatacaactgaagaattacttgaagcgttaaaaagc
aaaataatacaaccagaacaactcgtgacacattatagtaaactgagtgaaatcgaaagt
gcctatgatttattagaaatgcaacagatcataaagcgattaaattaatcatagagaat
15 gatatcacaatttaa

Sequence 3240

MLSDILPTGYEIGVLKGVKPGCTVAIVGAGPVGLAALLTAQFYSPSKIIMIDDDNRLE
TAKELGATHLINSKETETAIKKVKSLNPRGVDVAIEAVGIPQTFDLCQNLIGVDGTIANV
20 GVHGLPVQLDIDKLWIKNINVTGLVSGNTTEELLEALKSKIIQPEQLVTHYSKLSEIES
AYDLFRNATDHKAIKLIENDITI*

Sequence 3241

Contig_0743_pos_0_688

25 is similar to (with p-value 1.0e-50)

>sp:sp|P70814|RIBG_BACAM RIBOFLAVIN-SPECIFIC DEAMINASE (EC
3.5.4.-). >gp:gp|X95955|BARIBGENS_1 B.amyloliquefaciens ribB
, ribG, ribA, ribH & ribT genes. NID: g1592687.

atggatgatgctattcaactagcaaaaaatggtaaatggacaaacaggtgttaatccacca
30 gtaggatccgtgtgtgttaaaaacggtaggattgttaggttttaggtgcacatttaaaaaag
ggagataaaacatgccgaagtacaagctattgaaatggcagggtttaataaccgaaggtgct
accatatacgtttcattagaaccttgcacacaccatgggttcaacaccaccttgtgtgcat
aaaatcattgaagcgggcatacttaaggtcatctatgctgttaagatactacttttagta
agtaaggggtgacgagattctgagagaagctggtagaggttgaaattcaatataatgaa
35 aatgcagctgcattataccgtgacttttttactgctaaaagaaacgaagttccagaagta
actgtaaagggtctcatctagtctagatggtaacaagcaacagactttaatgaaagtaag
tggataacaacaagaagttaaagaagatgtttatcaattaagacatgagcatgatgca
gttattactcggcgctagaaaccattgaagcagacaatccattgtatacaaccagggttctt
gatggaaagcatccgattcgagttattcttttctaagaaaggtcaactcgattttaatcaa
40 caaatatttaagatactgcattCCCTG

Sequence 3242

MDDAIQLAKMVNGQTGVNPPVGSVVVKNRIVGLGAHLKKGDKHAEVQAIEMAGLNTQGA
TIYVSLEPCTHHGSTPPCVHKIIIEAGISKVIYAVKDTTLVSKGDEILREAGIEVEFYNE
45 NAAALYRDFFTAKRNEVPEVTVKVSSSLDGKQATDFNESKWITNKEVKEDVYQLRHEHDA
VITGRRTIEADNPLYTTRVPDGKHPPIRVILSKKGQLDFNQIFKDTASPX

Sequence 3243

Contig_0744_pos_2445_3218

50 is similar to (with p-value 5.0e-69)

>gp:gp|U96107|SCU96107_3 Staphylococcus carnosus N5,N10-met
hylenetetrahydromethanopterin reductase homolog, SceB precur
sor (sceB) and putative transmembrane protein genes, complet
e cds, and putative Na⁺/H⁺ antiporter NhaC (nhaC) gene, part
55 ial cds. NID: g2735503.

atgaaaaaaatcgctacagctacaattgcaactgcaggaatcgctacttttcgcatttgca
caccatgacgcacaagcagcagaacaaaataatgatgggtacaatccaaacgacccttat
tcatatagctacacttacacaatcgatgctgaaggtaactaccactacacttgaaaggt
aactggagtcagatcgtgtaatacttcatataactataataattataataactacaac

5 tactatgggttacaataactatagcaactacaataactacagtaattacaacaattacaac
aactatcaatcaaacacacgcaatcacaaagaacaactcaaccgactgggtggttaggc
gcaagctattcaacatcaagtagtaattgttcacggtacaacaacttctgcgccatcatca
aacgggtgatctttatcaaacgctcgtcagcatctggttaacttatacacttcagggtcaa
tgtacatatattatgtatttgacagagtaggtggcaaaatcgggttcaacgtggggtaaacgca
aacaactggggcaaacgctgcagcacgttctggtttacacagtaaacaaatctgcctgtaaa
gggtgcaattcttcaaacgtcacaaaggtgcatacggcacgtagcatcagttgaaggtgtga
aacagcaatgggttcaatcagagtttcagaaatgaactacggtcacgggtcaggtgtgtc
acttcacgtacaattctctgcgagccaagctgcttcataataactatattcactaa

10 Sequence 3244
MKKIATATIIATAGIATFAFAHHDAAEQNNDGYNPNDPYSYSYTYTIDAEGNYHYTWKG
NWSPDRVNTSYNNYNNYNNYGYNNYSNNYNNYSNNYNNYQSNNTQSQRTTQPTGGLG
ASYSTSSSNVHVTTSPASSNGVLSNARSASGNLYTSGQCTYYVFDVVGKIGSTWGNA
15 NNWANAARSGYTVNNSPAKGAIIQTSQGAYGHVAYVEGVNSNGSIRVSEMNHYGHGAGVV
TSRTISASQAASNYIH*

Sequence 3245
Contig_0744_pos_7402_7740
20 is similar to (with p-value 9.0e-30)
>sp:sp|P37279|ATCS_SYNPF CATION-TRANSPORTING ATPASE PACS (E
C 3.6.1.-). >pir:pir|S36741|S36741 cation-transporting ATPase
pacS - *Synechococcus* sp. >gp:gp|D16437|SYOPACS_1 *Synechoco*
ccus sp. DNA for PacS, complete cds. NID: g435124.
25 atggtagaaccaatcactgaatcgccactacttttttcaacaggtatagactcaccagtt
agcatggattcgtaatagaagtatcaccttagtgactttgccatctacaggatatctt
tcgccgggttttattagtaaagtatctccgactttaactttatcaagtggaagcataatt
tctttattttctttaattactcgctgtcttcttcgctgttaaatttaacaattcgctta
gcatggtagtcgtgtgattggcacgtgtcttcaaatatttaccagaagaattaacgta
30 attaaaaatagcactttgtttcaaaaataaatgcggtga

Sequence 3246
MVEPITESPLLFSTGIDSPVSMDSIEVSPVLTPSTGIFSPGFISKVSPTLTSSGSI
SLFSLITRASFAKFMNSLNALVCDLARASKYLPRRINVIKIALVSKYKCG*

Sequence 3247
Contig_0744_pos_9912_8887
is similar to (with p-value 3.0e-66)
>sp:sp|P56157|DP3A_HELPY DNA POLYMERASE III, ALPHA CHAIN (E
C 2.7.7.7). >gp:gp|AE000646|HPAE000646_3 Helicobacter pylori
section 124 of 134 of the complete genome. NID: g2314635.
atgggatttgatgaaacgagtttaaatgagatttcaaaacttattccacataaattaggt
ataactcttgaaagaagcataccaaaagccagagtttaagcatttgttcacgtaatcat
agaaatgaacgttggtttgaaagtgaagtaaaagtttagagggattaccaagacatacgtct
acgcattgctgcagggtatcattatcaatgatcaaccattattcaaatttggccccattaa
actggtgatacaggatttaatacgcagtggaactatgcagagcggaacgtaaggatta
ttaaaaattgatttcttgggattacgcaactctatcaattattcattcaaattattttaca
gttaaaaaggattttaaataataatattgatatagaagctataccttatgatgataaaaaa
gttttttgatttattatcaaacggtgacactacagggtatatttcaattggaatcagacggt
gttagaagcgatttaaaaagattgcaacccgaacattttgaagatatcgtagctgtcaca
tcattatatagaccaggaccaattggaagaaataccaacttatataaccgtagacataat
cctaaccaatttgcattatttaccatcgagtttagaaccaactttaaaaaaacatcattggt
gttatcatttatcaagaacaaataatgctgaatagaagtcgaagtcaagttgctggttttagttat
gggtgaagcagatattttaagaagggcaatgagtaaaaagaatcgtgcaactcttagaaagt
gagcgtcaacatttcattgatggtgcaaaaaataacggttacgatgaacagataagtaag
caaatttttgatttaataacttaagtttgcagattatgggttcccacgtgcccatgctgtt
agttactcaaaaattgcatacatttatgagctattttaaagtgccacaaatcagtacgaa
ccggttatgcctggcttcaaatatgaagcaccaccaaatcaaaataaaatcaatccttatg
aattaa

Sequence 3248

MGFDETSLSNEISKLIHPKLGITL EAYQKPEFKAFVHRNHRNERWFEVSKKLEGLPRHTS
 5 THAAGIIINDQPLFKFAPLTTGDTGLLTQWTMTAEERIGLLKIDFLGLRNL SIIHQIILQ
 VKKDLNINIDIEAIPYDDKKVFDLLSNGD TTGIFQLES DGVR SVLKRLQPEHFEDIVAVT
 SLYRPGPMEEIPT YITRRHNP NQFAYLHPDLEPILKNTYGVIIYQE QIMLIASQVAGFSY
 GEADILRRAMS KKNRAILES ERQH FIDGAKNNGYDEQISKQIFDLILKFADYGF PRAHAV
 SYSKIAYIMS YLKV PQISTEPLCLASNMKHHQIKIKSILMN*

10 Sequence 3249

Contig_0746_pos_2558_3991

is similar to (with p-value 2.0e-45)

>sp:sp|P26648|SUF1_ECOLI SUF1 PROTEIN PRECURSOR. >gp:gp|U28
 377|ECU28377_116 Escherichia coli K-12 genome; approximately
 15 65 to 68 minutes. NID: g882431. >gp:gp|AE000384|AE000384_1
 Escherichia coli K-12 MG1655 section 274 of 400 of the compl
 ete genome. NID: g2367186.

atgtataataaagtttttgcatttttaattataattttttccataataattattgcgtct
 aatgatacttttcgcagaaagtaagaatgatatgatgaatatgaaagaagataagaaaaat
 20 acaatggatattgacaaatattgaaacatcatgacgaaagaagaaattaaattcttcacaa
 ggaaaaaatgaaataatatttctaaagttgcagagtcaaaaaaagataacaatggttat
 aaaaattatacattaaaagctcaggaaggaaagacagagttttacaaaaataatttttct
 aatactctaggtacaatggaaatttacttggaccaactttaaaattaaaaaaaggagat
 aaagttaaaaattaagttataaataactttagatgaaataacaacatttcattggcatgga
 25 ttagaagtaaaaggaaagtggaaggccttctcagggtataaaaaccaggaaaagaa
 aaaactataaaaatttgagggttaatacaagattctgctacgttatgggtatcacccccacccc
 tctccaaatacagctaaacaagtttataatggccttatcaggattattatatatagaagat
 agtaaaaaagaataattatcctagtgattatggaaaaaatgatttgcctataataatccaa
 gataaaacatttgtatctaaaaaattaaattattcaaaaacgaaagacgaagatggcact
 30 caaggtgatactgttcttgtgaacggaatagtaaaaccccaactgacaacaaaagaagag
 aaatatcgtttgagacttttaaatgggttctaattgctcgagatttaaatcttaagctaagt
 aataatcaaaagtttttgagtatatgttctcagatggcgggtcaattaaaaaacgctaaaaaa
 ttaaaagaaattaatttagctccttcagaaagaaaagaaatagtaataagatttatctaaa
 atgaaaggcgagaaaatcagctctgggtgataatgataaaactgtaattttaccgattagt
 35 acaaaagagaaaagttctaacaagggaatacaccaaaaagtaagtaaaaaataaaaatta
 gaaggtatgaatgataatgttaccattaatggtaataaattcgatcctaaaagaatagat
 ttacacaaaaagttaaaccagaaagaagtatgggaaattgaaaacgtcaaagataaaaatg
 ggtgggtatgaaacatcctttccacatccatggaacgcaatttaaaagttttatctgtggat
 ggggagaaaccacaaaagatatgaggggtaaaaaagatgttatatctttggaacctgga
 40 caaaaagctaaaatagaggttgtatttaaaaataactggaacatacatgtttcactgtcat
 atacttgagcatgaagataatggaatgatgggtcaaataaaagtaacaaactaa

Sequence 3250

MYNKVFAILIIIFSIIIIASNDTFAESKNDMMNMKEDKKN TMDMTNMKHHDERKKLNSSQ
 45 GKNEIIFPKVAESKKNNGYKNYTLKAQEGKTEFYKNNFSNTLGYNGNLLGPTLKLKKG
 KVKIKLINLNDENTTFHWHGLEVNGKVDGGPSQVIKPGKEKTIKFEVNQDSATLWYHPHP
 SPNTAKQVYNGLSGLLYIEDSKKNYPSPDYGKNDLP IIIQDKTFVSKKLNYSKTKDEDGT
 QGDTVLVNGIVNPKLT TKEEKIRLRLNLSNARDLNLKLSNNQSFEYIASDGGQLKNAKK
 LKEINLAPSERKEIVIDLSKMKGKISLVDNDKT VILPISNKEKSSNKGNTPKVSKKI KL
 50 EGMNDNVTINGNKFDPKRIDFTQKLNQKEVWEIENVKDKMGMKHPFHIHGTQFKVLSVD
 GEKPPKDMRGKKDVISLEPGQKAKIEVVFKNTGTYMFHCHILEHEDNGMMGQIKVTN*

Sequence 3251

Contig_0748_pos_3358_4230

55 is similar to (with p-value 1.0e-38)

>sp:sp|P44948|FPG_HAEIN FORMAMIDOPYRIMIDINE-DNA GLYCOSYLASE
 (EC 3.2.2.23) (FAPY-DNA GLYCOSYLASE). >pir:pir|A64104|A6410
 4 formamidopyrimidine-DNA glycosylase (fpg) homolog - Haemop
 hilus influenzae (strain Rd KW20) >gp:gp|U32776|U32776_1 Hae

mophilus influenzae Rd section 91 of 163 of the complete genome. NID: g1573969.

atgcctgaactacctgaagttgaacatgttaaaagaggtattgagccatttataaaaagt
gcaaaaatagagaaagtaacttttgcataaaaatgtaattaacggtaagaataataaccgt
5 gagacgattataaaaaggtatggaattagatacttttaaaaaacttactgaagggtatggt
ataaaaaaagttgaaagaagaagtaagtacattatttttatatagcggatcatgacgat
gatagaatccttagttagtcatttaggtatggcaggcggattccttggttgtaataacctt
gatgagataagtaacccgaattatcgaaagcattggcaagtcattttcgatttggataat
aaacaaaaatttagtctattctgatatcagacgggttggagaaattagaaatatagtcaat
10 ttgatagttatccatcctttattagaaatcgctccagaaccatttgaagaggtggcattt
gaacactatttgaatgtttgacaatgaaaaaatataagaataaaccaataaaacaaacg
attccttgatcatcgtgttatagcaggagctggaaatatctatgcctgtgaagctttattc
agagctgggtattactccggataaaaattactaattcactcactaaacaagaaagaaaatcc
ctctttttattatgttcgagaagttttagaagaggggtataaaaatatggaggtactagtatt
15 tcagattatagggcatgcagatggtaaaactggacaaatgcaattacatttaaatgtatat
aaacaaaaaaagtgaaggtttgtggtcattcgattgaaacaaaagtgatagctggtaga
aatagtcatttttgcocaaactgtcagagataa

Sequence 3252

20 MPPELPEVEHVKRGIEPFIKSAKIEKVTFKKNVINGKNNNRETIIKGMELDTFKKLTEGYV
IKKVERRSKYIIFYIADHDDDRILVSHLGMAGGFFVNNLDEISTPNYRKHWQVIFDLN
KQKLVS DIRRFGEIRNIVNFDSPSLLEIAPEPFEEVAFEHYLECLTMKKYKNKPIKQT
ILDHRVIAGAGNIYACEALFRAGITPDKITNSLTKQERKSLFYVREVLEEGIKYGGTSI
SDYRHADGKTGQMLHLNVYKQKKCKVCGHSIETKVIAGRNSHFCPNCQR*

25

Sequence 3253

Contig_0748_pos_4883_5752

is similar to (with p-value 4.0e-89)

30 >gp:gp|Z82038|CTZ82038_4 C.thermosaccharolyticum etfB, etfA
, hbd, thlA and actA genes. NID: g1667352. >gp:gp|Z92974|TTB
CSOPRN_6 T.thermosaccharolyticum BCS operon DNA. NID: g19033
26.

gtgtttgggtgtgtatttaaggatatacctgcctatgaactaggtgcaacagttattcgt
caaattttagaacatagtc aaatagatccaatgaaatcaatgaagttattctaggaac
35 gtattacaggcagggtcaaggacaaaatcctgctcgtattgctgcatcatgggtgtg
ccagaagcgggtaccttctttactgtaaataaagtttgcgggttctggattaaaagcgatt
caacttgcttatcaatctattgtagcgggagataatgagattgttatcgtcggaggcatg
gaaagtatgtctcaatctccaatgcttcttaaaaatagtcgttttcggttttaaaatggga
aatcaaacctttagaagatagtagatgatagctgatggtttaactgataagtttaattgattac
40 catatgggtatcacagccgaaaatctagttgaacagtatcagatttagtcgtaagaacaa
gatcaatttgcattcgattctcaacaaaaagcatcacgtgcacaaagctgggtgtattt
gatgctgaaattgtacctgtagaggtaccacaacgtaaaggcgaccccttaattttct
caagatgaaggcatttagacctcaaacgacaattgataagttagcacaactccgtccagca
tttaaaaaagatggatcagtaactgctggtaatgcatccgggtatcaatgacgggtgctgct
45 gctatgctcgttatgacggaggacaaagcgaagcattgggcttacaacctatagctgta
ttagatagttttgggtgcgagtggtgtggcgcttcaattatgggtattcgacgcacaaaa
acaaaccaaagaagaatagcaactgttaa

Sequence 3254

50 VFGGVFKDIPAYELGATVIRQILEHSQIDPNEINEVILGNVLQAGQGQNPARIAAIHGGV
PEAVPSFTVNKVCGLKAIQLAYQSIVAGDNEIVIAGGMESMSQSPMLLKNSRFGFKMG
NQTLEDSMIADGLTDKFN DYHMGITAENLVEQYQISRKEQDQFAFDSQKASRAQQAGVF
DAEIVPVVEVPQRKGDPLIISQDEGIRPQTITDKLAQLRPAFKKDGSVTAGNASGINDGAA
AMLVMTEDKAKALGLQPIAVLDSFGASGVAPSIMGIRRTKTNRNSNC*

55

Sequence 3255

Contig_0753_pos_1053_58

is similar to (with p-value 3.0e-40)

>gp:gp|AB011003|AB011003_1 Candida albicans CaUP1 gene for

UDP-N-acetylglucosamine pyrophosphorylase, complete cds. NI
D: g3273313.

atgttagacaaaaatcaattggaaaagtataaccaagagcatttgtatgaatatgaaaaa
ttaatgagtagtaaatgaaaagaatgcttttagatgaaaaagtagatcagttaaatcttgca
5 gaaattcaagatttatatcaagatttatatgttaatagaaaaactattgatgatgtatct
tctgtatctgaagtcaaatatgaagtgaatcacgactcaatgaagaagaacgacataca
tatgaacaaaaagggttatgaggcaatacgaatgggtgaatttgctgtattattgatggct
ggaggacaagggtacgcgttttaggatataaagggcctaaagggtcttttgaaatagagggt
acgagtttatttgaacttcaggcgctcaactgattcgtttaaaagaagaacggccac
10 acaatttaattggatatattatgacaagtgcatttaatacataaagatacaatagagtatttt
aaacaacataaataattttaactatgatgccaatcatattcatttctttaagcaagataac
attgttgctttaagtgaagaaggaaagcttggttttaaatagagatggacataataatggaa
acacctaattggtaattgggggtgattcaagtctcttaagaaagcaggataccttgataag
atgcaacaagatcacgtcaaatatattcttcttaataacattgataatgtcttagttaaa
15 gtttttagaccggtatttggcgggtttacagtgacacaaagtaaaagacatcacatcaaaa
acaattcaacctaaagatagtgaagtgtaggtcggttgtaaatgttgattgtaaagac
actgtgttagagtattctgaattaatgtcatatttccaggagctgcactccagttcattt
ctgaaactaaaaatacttccatcgctattcacgcgctctacaaacgctacgtgaccatagt
aaccagcgtcagtttgtgcaattgagcctactgtag
20

Sequence 3256

MLDKNQLEKYNQEHLYEYKLMSSNEKNALDEKVDQLNLAEIQDLYQDLYVNRKTIDDVS
SVSEVKYEVKSRLNEEERHTYEQKGYEAIIRNGEFAVLLMAGGQGTRLGYKGPKGSFEIEG
TSLFELQARQLIRLKEETGHTINWYIMTSDINHKTIEYFKQHKYFNVDANHIHFFKQDN
25 IVALSEEGKLVLRDGHIMETPNGNGGVFKSLKKAGYLDKMQQDHVKYIFLNNIDNVLVK
VLDPLFAGFTVTQSKDITSKTIQPKDSESVGRLVNVDCKDTVLEYSELMSYFQELHSSSF
LKLKYFHRYSRALQTLRDHSNQRFVQLSL*

Sequence 3257

Contig_0754_pos_7866_0
is similar to (with p-value 4.0e-44)
>sp:sp|P36649|YACK_ECOLI PROBABLE 53.4 KD BLUE-COPPER PROTE
IN YACQ PRECURSOR. >gp:gp|AE000121|AE000121_8 Escherichia co
li K-12 MG1655 section 11 of 400 of the complete genome. NID
35 : g1786306.
atgtataataaaagtttttgcaatttttaattataattttttccataataattattgcgtct
aatgatactttcgcagaaagtaagaatgatatgatgaatatgaaagaagataagaaaaat
acaatggatatgacaaatataaacatcatgacgaaagaagaattaaattcttcacaa
ggaaaaaatgaaataatatttcttaagttgcagagtcaaaaaaagataacaatgggtat
40 aaaaattatacattaaaagctcaggaaggaaagacagagttttacaaaaataatttttct
aatactctaggctacaatggaaatttacttggaccaactttaaaattaaaaaaaggagat
aaagttaaaattaagttaataataaacttagatgaaaatacaacatttcattggcatgga
ttagaagtaaaatggaaaagtggatggagggccttctcaggttataaaaccaggaaaaagaa
aaaactataaaaatttgaggttaatacaagattctgctacgttatgggtatcacccccacccc
45 tctccaaatacacgctaaacaagtttataatggcttatcaggattattatatatagaagat
agtaaaaaagaataattatccttagtgatttatggaaaaaatgatttgcctataataatccaa
gataaaacatttgtatctaaaaaattaaattattcaaaaacgaaagacgaagatggcact
caaggtgatactgttcttgtgaacggaatagtaaaccccaactgacaacaaaagaagag
aaaatacgtttgagacttttaaatgggttctaattgctcgagatttaaatcttaagctaagt
50 aataatcaaaagttttgagtatttgccttcagatggcgggtcaattaaaaaacgctaaaaaa
ttaaataagaataatttaggtccttcagaaagaaaagaataagtaattatcttaaa
atgaaaggcgagaaaatcagttctgggttgataatgataaaactgtaattttaccgattagt
aacaagagaaaaagttctaacaaggtaatacaccaaaagtaagtaaaaaataaaatta
gaagggtatgaatgataatgttaccattaatggtaataaattcgatcctaaaagaatagat
55 ttacacaaaagttaaaccagaaagaagtatgggaaattgaaaacgtcaaagataaaaatg
gggtgtatgaaacatcctttccacatccat

Sequence 3258

MYNKVFALIIIFSIIIIASNDTFAESKNDMMNMKEDKKNMTMDMTNMKHHDERKKLNSSQ

GKNEIIFPKVAESKKDNNGYKNYTLKAQEGKTEFYKNNFSNTLGYNGLLGPTLKLKKGD
 KVKIKLINNLDENTTFHWHGLEVNGKVDGGPSQVIKPGKEKTIKFEVNQDSATLWYHPHP
 SPNTAKQVYNGLSGLLYIEDSKNNYPSDYGKNDLPPIIQDKTFVSKKLNYSKTKDEDGT
 QGDTVLVNGIVNPKLT'TKEEKIRLRLNLSNARDLNLKLSNNQSFEYIASDGGQLKNAKK
 5 LKEINLAPSERKEIVIDLSKMKGEKISLVDNDKTVILPISNKEKSSNKGNTPKVSKKIKL
 EGMNDNVTINGNKFDPKRIDFTQKLNQKEVWEIENVKDKMGGMKHPFHIH

Sequence 3259

Contig_0755_pos_2032_3393

10 is similar to (with p-value 8.0e-68)

>gp:gp|L41217|RICNRAMP_1 Oryza Sativa integral membrane pro
 tein (OsNramp) mRNA. NID: g2231131. >gp:gp|S81897|S81897_1 O
 sNramp1=Nramp1 homolog/Bcg product homolog [Oryza sativa, in
 dica, cv. IR 36, etiolated shoots, mRNA, 1967 nt]. NID: g147
 15 0319.

atgggggtgattttattgaattccaataataacaatcatgaacaacaacgaagtttagat
 gaaatcaataacaccataaacttcaatcataatgatagtgcaagtcacaaatcttctggct
 ttttaggacccgggattgcttgttgagttggttacatggatcctggaaattggattaca
 tccatgcaaggaggagcacaatatggctatacctgttattcataatcttaactcttagc
 20 ttatctgctatgctgttacaaagtatgactgtgagattaggaatagcaactggatggat
 ttagcacaatgacacgtcatttttaataagcctgtagcaattatgttctggattatt
 gcagaattagcaattatcgctactgatattgcagaagttataggtagcgctatcgatta
 gatttaactcttcggcataccatttaattgtaggcgcatataactcgtatttgatgtattt
 ttattattattcatcatgaaatttggcttttagaaagattgaagctatcgtgggaacgta
 25 atctttaccgtattggccattttgtatttgaagtttatatttcttccacaaataaca
 gatattgcttaattggttttgtgctcataaagaaattattacaaaccaagggatactttat
 attgcaactaggtatcataggtgctactattatgccacataacttatattacattcttct
 attgtacaatctcgaaaatatgatagacacagttatcatgaaaaagcacaagcgattaag
 tatgctactatcgactcctaataacagctatccatcgctttttagtcaattgcttatta
 30 cttacacttggtgcagcgctatttttggaaactaaaactgaagatttggtgggtttttat
 gatctttattttggctctaaaaacagaacctgcttttaggtgcaacgcttggtggtattatg
 agtactttatttgcgtgttgcccttttagcttctggtcaaaattcaactataacgggaacg
 tttagcaggccaaattgtgatggaaggatttcttaaaattatccattccaaattggttacgt
 cgtcttatcactcgggtctttagcagtgatacctgttatcatttgccttatagtatttaaa
 35 ggaaatactgaaaaaattgaacaattacttgccttttctcaagtggtcttgagtattgct
 ttgccattttcggttaataaccgcttcaattagctacaagtaatacaaatcttatgggtcct
 ttaagaataaaaacatggattaacatcatttcttggttactcataattgtcttaagtggga
 cttaacgtatatcttatcattcaaacattccaagaattatga

40 Sequence 3260

MGVILLNSNNNNHEQQRSLDEINNTINFNHNDASQKFLAFLGPGLLVAVGYMDPGNWIT
 SMQGGAYGYTLLFIILISSLSAMLLQSM TVRLGIATGMDLAQMTRHFLNKPVAIMFWII
 AELAIATDIAEVIGSAIALDLIFGIPLIVGALITVFDVFLLLFIMKFGFRKIEAIVGTL
 IFTVLAIFVFEVYISSPQITDMLNGFVPHKEIITNQGILYIALGIIGATIMPHNLYLHSS
 45 IVQSRKYDRHSIHEKAQAIKYATIDSNIQLSIAFVNVCLLLTLGAALFFGKTEDLGIFY
 DLYLALKTEPALGATLGGIMSTLFAVALLASGQNSTITGTLAQIVMEGFLKLSIPNWL
 RLITRSLAIVPVIICLIVFKGNTEKIEQLLVFSQVFLSIALPFSLIPLQLATSNQNLMP
 FKNKTWINIISWLLIIVLSGLNVYLIIQTFQEL*

50 Sequence 3261

Contig_0755_pos_6203_5673

No hits found

atgggattaaataaattagggatagtccttaatgacaaaaatccgacccatattcatcaa
 gcagaaaaagatttgcataatttagcacctcaagttagagggatagtaggcgacgaaatt
 55 actatgatgcttcaacaaaacgaaggacatgttgagtagtttgagtggtgctgca
 ccacttgtagcaggaaaatactcgttataattacgtgatacctaaagaaggctctaacta
 tgggttgataatattggtgatacctaaaactgcacaaaataaagaaggctgcgtataagttt
 atgaatttcttacttagacgcgcaaaatagtgccagaatacggaaatgggttaggatgca
 acacctaataaagcagctcgaagtaagttgcctaaaaggtaagaatgattatagattt

tatccatcaaatcaagaacagcaacgggttagaagtctataaagatttaggtcaaacgtct
ctcagtgaatataatgaaagcttttttaaattttaaaatgtcttttaaaatag

Sequence 3262

- 5 MGLNKLGYSLNDKNPETHIQAEKDLHNLAPQVRGIVGDEITMMLQQNEGHVAVVWSGVAA
PLVQENTRYNYVIPKEGSNLWFDNMVIPKTAQNKEGAYKFMNFLDAQNSAQNTIEWVGYA
TPNKAARSKLPKKVRNDYRFYPSNQEQQRLEVYKDLGQTSLSSEYNESFLNFKMSLK*

Sequence 3263

- 10 Contig_0755_pos_3989_3612
is similar to (with p-value 8.0e-25)
>sp:sp|P23861|POTD_ECOLI SPERMIDINE/PUTRESCINE-BINDING PERI
PLASMIC PROTEIN PRECURSOR (SPBP). >pir:pir|D40840|D40840 spe
rmidine/putrescine transport protein D - Escherichia coli >g
15 p:gp|D90747|D90747_2 Escherichia coli genomic DNA.(25.1 - 25
.5 min). NID: g1651548. >gp:gp|AE000212|AE000212_9 Escherich
ia coli K-12 MG1655 section 102 of 400 of the complete genom
e. NID: g1787358. >gp:gp|M64519|ECOPOTABCD_4 E.coli transpor
t protein (potA, potB, potC and potD) genes, complete cds. N
20 ID: g147325.
atgtcatctgaaatggttttagtttctcttattgctgatttagtcatttttaggtttttca
acgctatttaggttttgttggtacaatggtaattgaaggtagaaaaaaccttgctgctagt
ttactcatagcagcggctatcgtaggtttatttacgactaatttaacgcaatgggttta
tggatgattgctgctgattagactttttgcaaaaaagataaaacagatgtaaatgaaaa
25 tctacggctcaacttcgtcaaaaccattcaaaagagccaaagtgttggaatcatcaaca
aaccaacaacagaaagatgcttgggacacctgaacaagaaatcaacaacaaaaaaaggac
gatccatatatatattaa

Sequence 3264

- 30 MSSEMFLVSLIASLVILGFSTLLGFVGTVMVIEGRKNLAASLLIAAAIVGLFTTNLIAMVL
WMIAAIRLFAKKDKTDVNENATAQLRQNHKSQSQDWNHQNNQQQKDAWDPEQEINKQKKD
DPYIY*

Sequence 3265

- 35 Contig_0756_pos_5439_6113
is similar to (with p-value 3.0e-39)
>gp:gp|U67196|TMU67196_1 Thermotoga maritima DNA-binding re
sponse regulator (drpA) and histidine protein kinase (hpkA)
genes, complete cds, thymidine/pyrimidine phosphorylase homol
40 og gene, partial cds. complete cds. NID: g1575576.
atgattaattgcttaatcgtagacgatgataaaaagttattgcaatatgtttcaagtcatt
ttagaaagagaaagtattcaaacacatactttcacaagtggagaagcatcactagatttt
cttgaataaaaaatgttgatattgcatagtagatattatgatgagtggaaatggatggt
tttgagctttgtcagactttgaaagatgattatcatattcctgtcataatgttaacagct
45 agagatgcattaagtataaaagaacgtgcattttctaagtgggactgacgattatgtcact
aaaccttttgagggttaagaattattatttagaattaaagctgtcttaagacgatatcaa
attaatgctgataacgagttacaacttggcaacttaattataaatcagtccttacatggaa
attactgtgggttcaaaaacgatgaatcttccaaacaaagaatttcagttgttattttta
ttagcctctaactctaacaatattttcactcgagatgatattattgaaaaaatttggggc
50 ttcgattatgaaggagatgatcgactgttgatgttcatattaaaagattacgtcaacgt
ttatctaaattgaaatcatcagtatcaattcaaaactgtaagaggtcaaggatatagggtg
gacaaaaatggttaa

Sequence 3266

- 55 MINCLIVDDDKKLLQYVSSHLERESIQTHFTSGEASLDFLENKNVDIAIVDIMMSGMDG
FELCQTLKDDYHIPVIMLTARDALSDKERAFLSGTDYVTKPFVKEKLLFRIKAVLRRYQ
INADNELQLGNLILNQSYMEITVGSKTMNLPNKEFQLFLLASNPKHIFTRDDIIEKIWG
FDYEGDDRTVDVHIKRLRQLSKLSSVSIQTVRGQGYRVDQNV*

Sequence 3267

Contig_0757_pos_1852_2325

is similar to (with p-value 1.0e-26)

5 >gp:gp|AJ000084|PMAJ84_3 Proteus mirabilis ccm and pat gene
s and partial ygbA gene. NID: g3395515.
atgaatgctaatagccccagtagatagcttccctagtttacttagttatcgctggttct
gagatagacatggaaaatagatagcacaatacaagagagtgccctaggtttatttgct
gtagagttgaaagagacgggagaatgggttaggtttataggtgtgaattacgtttcaaaa
gatagccattacccttttaaaagagttaccggtttatgaaataggttgagggttaattcca
10 gaagtatggggaaatgggtctagctacagaaggggcagaagccgtaataaataatgctaga
gataaaggaaataaagaattatagttttacttctgaaaataatttgccttctagaaaa
gtcatggaaaaattaggaatgacttttttagacaattttgaatatccgaatcttagtaaa
taccatccccctaaacgtcatgtaagatattataaagagctacttctcttctga

15 Sequence 3268

MNANAQVRRYFSPILLSYRRSEIDMENMDSIIQESGLGLFAVELKETGEWLGFIVNYVSK
DSHYPFKELPFYEIGWRLIPEVWGNGLATEGAEAVMKYARDKGIKELYSFTSENNLPSRK
VMEKLGMTFLDNFEYPNLKYHPLKRHVRYKELLPS*

20 Sequence 3269

Contig_0757_pos_4021_3656

is similar to (with p-value 2.0e-29)

25 >sp:sp|P46378|FAS6_RHOFA HYPOTHETICAL 21.1 KD PROTEIN IN FA
SCIATION LOCUS (ORF6). >pir:pir|F55578|F55578 hypothetical p
rotein 2 (ipt 3' region) - Rhodococcus fascians plasmid pFiD
188 >gp:gp|Z29635|RFCCIPTFD_6 R.fascians (D188) genes for P4
50 cytochrome, isopentenyltransferase and ferredoxine. NID:
g455000.
atggacaaaaagcttcaaaagaatattgagaaacgtcataaagaagaacaaaaacaacgt
30 gaagctaatacagaacaacgtattaaagacatgaaaaaactcaaaaatacgaagagcaa
gtggcttaactccttggttaaaatagatcacgaaattgagaaaaaaggcgaaaaactagaa
aaagataatcgtaaaagatattaaaaaattagataaaaagcttcaaaagaatattgaaaaa
cgtcataaagaagaacaaaaacagcgtgaagaagcagagaaagctagaaaaaagaattt
aaaaaatatgaaaattacgtggctgacagtgctgtaaaacaacataaggaatcaaatcat
35 tcttaa

Sequence 3270

MDKKLQKNIEKRHKKEEQKQREANQKQRIKDMKKTQKYEEQVGLTPGKIDHEIEKKGEKLE
40 KDNRKDIKKLDKKLQKNIEKRHKKEEQKQREEAEKARKKEFKKYENYVADSVVKQHKESNH
S*

Sequence 3271

Contig_0757_pos_3193_2627

No hits found

45 atgaaaagaattgctgtttattgttggtgcaagtaaaagggaacccatcttatgtttaa
gaggcatacgaattaggaaaatatatggctgaacaaggatacagagcttgattcggagca
ggatcagtcggcattatgggagctattcaagatggcatacttgagcatggcggtaaagct
atcgggtgcatgcctaaaaatgttagatgaacgagaaataacaagccaaaaagtaagtga
cttatattagtagattctatgcatgaacgaaaaataaaatgactgaacttgccgatgct
50 tttattatggctccaggcggtgctggttcattagaagaattttttgaaatgtatagttgg
tctcaaatgggtatacaccaaaagcctataggtgtatttaatttaaatggattctttgag
ccactacaacacttaatcgaccatatgattaaagaaggatttattgatgagaaatatcaa
aagcttgacctttatatgatactaaagaatcactcatcgaaggacttaaacattacaaa
ccacttggtgtacgtacatacattaa

55

Sequence 3272

MKRIAVYCGASKGNPSYVKEAYELGKYMAEQGYELVFGAGSVGIMGAIQDGIIEHGGA
IGVMPKMLDEREITSQKVSELILVDSMHERKNKMTTELADAFIMAPGAGSLEEFFEMYSW
AQIGIHQKPIGVFNNGFFEPLQHLIDHMIKEGFIDEKYQKLAPLYDTKESLIEGLKHYK

PLGVRTYD*

Sequence 3273

Contig_0759_pos_3773_3006

5 is similar to (with p-value 9.0e-58)

>gp:gp|U38892|SSU38892_2 Synechocystis sp. ruvB gene, complete cds, and secA gene, partial cds. NID: g1256587.

atgttttaaaataggaaatttagaattacaatctcgtttacttttaggtactggaaaattt
gaaaatgaagaggttcagtcaaaagcaattgaggcatctgaaacaaatgtacttacattt
10 gcagtaagacgtatgaatttatatgatcgtaacctacctaaccacttgcaaacgttaat
ttaaagaattttatcacttttccaaatactgcaggtgccaaaacagctcaagaagctatc
agaattgctgaaaattgctaatacgcgcaggtgtatgtgacatgattaaaagtcgaagtcatt
ggatgatgacgaaacattattacctgatccattcgaaacatacaggcatgcaaagtattg
ttagaaaaagggttacactgtttgtccttacatctctaacgatttagtttagctcaacgt
15 ttagaagaattgggtgtacacgcagttatgccacttgcacccctattggtacaggaaga
ggatattaataacccattaaatttaagttatattatcgaaaatgctagtgtacctgtaac
gtagatgctggtattgggtccctaaagatgcgtgtcatgccatggagcttggcgagat
ggatattttactcaacacagcatttcagcggcaaaagatcctgtgaaaatggctgaagca
atgaaattaggtataaatgctggcagactttcataatgaagctggacgcattcctgttaag
20 tatactgcacaagcatctagtccatcagaaggtttagggttcttgttaa

Sequence 3274

MFKIGNLELQSRLLLGTKFENEVQSKAIEASETNVLTFAVRRMNLYDRNLPNPLANVN
LKDFITFPNTAGAKTAQEAIIRIAEIANHAGVCDMIKVEVIGDDETLPLDPFETYEACKVL
25 LEKGYTVCPYISNDLVLAQRLEELGVHAVMPLASPIGTGRGINNPLNLSYIIENASVPVI
VDAGIGSPKDACHAMELGADGILLNTAISAADKDPVKMAEAMKLGINAGRLSYEAGRIPVK
YTAQASSPSEGLGFL*

Sequence 3275

30 Contig_0763_pos_1317_2216

is similar to (with p-value 1.0e-47)

>gp:gp|U76260|PAU76260_2 Peptostreptococcus asaccharolyticus
alpha- and beta-subunits of L-serine dehydratase (sdhB) and
d (sdhA) genes, complete cds. NID: g2315864.

atgtttgattcaattagagagacaatagattattcagttgaaaataacatcagttttgct
gatatgatgattaatgatgaaatggaaagagaaggttaaattctcgcaagaagtgcgtgat
ttaatgagacaaaacttaaatgttatgcgtgaggcagttgaaaaaggtagcactggcgat
gggtgtggaagtgtaacaggttatacaggtcatgatgctgctaaaacttcgagattacaac
gagaataatcatgcattatcaggtcatgaaatgattgatgcagtcgaagggtgcagttgca
40 acgaacgaagtcaatgcagcaatgggtattatttgtgctactccgacagctgggttcctcg
ggaacgattccccggcgtaattttaaattagaaaaactcataatatcactgaagatcaa
atgatagattttctattcacatcagctttattcggcgagtagttgcaaacaacgcgagc
gttgcgggtgcaactgggtgtgtcaagccgaagtgggttcggcatctgcaatggctgca
gctgctgcagtatcaatttttaacgggtcaccagaacaatcaggacatgccatggcattg
45 gcaattagtaacttattaggcttagtttgcgatccagttgctgggttagttgaaattcca
tgtgtaattgcgtaattgctattgggttcaggaaatgcattaatatctgcagaccttgacta
gctggagttgaaagtcaaattccagttgatgaagtcataaggtgctatggatagagtaggt
cgtaatttacctgcatcattaagagaaacaggttttaggcggttttagcaggtacacctact
ggcgaagaaattaaacgtaaaatattcggcggaagcagacaacatggttaaaaataaataa
50

Sequence 3276

MFDSIRETIDYSVENNISFADMMINDEMEREKGSREEVRDLMRQNLNVMREAVEKGTGTD
GVESVTGYTGHDAAKLRDYNNHLSGHEMIDAVKGAVATNEVNAAMGIICATPTAGSS
GTIPGVIFKLEKTHNITEDQIDFLFTSALFGRVVANNASVAGATGGCQAEVGSASAMAA
55 AAASVIFNGSPEQSGHAMALAI SNLLGLVCDPVAGLVEIPCVMRNAIGSGNALISADLAL
AGVESQIPVDEVIGAMDRVGRNLPASLRETGLGLAGTPTGEEIKRKIFGEADNMVKNK*

Sequence 3277

Contig_0765_pos_2844_3362

is similar to (with p-value 3.0e-17)

>sp:sp|P04928|SANT_PLAFN S-ANTIGEN PROTEIN PRECURSOR. >pir:
 5 pir|B22011|YAZQN7 S-antigen precursor - Plasmodium falciparum (strain NF7/Ghana) >gp:gp|M10130|PFASA7_1 Plasmodium falciparum (isolate NF7) S antigen gene, complete cds. NID: g160670.
 gtgcttctgcattgcctccttcttctgccttcggttggtcagattgaggtgcttctgcat
 10 tgcttccttcttctgtcttcggttggtcagattgagctgcttctgcattgcctccttctt
 ctgccttcggttggtcagattgagctgcttttgcttgcttccttcttctgtcttcggtg
 gctcagattgaggtgcttctgcattgcttccttcttctgccttcggttggtcagattgag
 ctgcttttacggtgcttccttcttctgtcttcggttggtcagattgagctgcttctgcat
 tgcctccttcttctgccttcggttggtcagattgagctgcttctgcattgcctccttctt
 ctgccttcggttggtcagattgaggtgcttctgcattgcctccttcttctgccttcggtg
 15 gctcagattgagctgcttctgcattgcctccttcttctgccttcggttggtcagattgag
 gtgcttctgcattgcctccttcttctgcttgcaatgtag

Sequence 3278

VLLHCLLLLPSLAQIEVLLHCFLLSSSLAQIELLLHCLLLLPSLAQIELLLRCFLLSSSL
 20 AQIEVLLHCFLLPSLAQIELLLRCFLLSSSLAQIELLLHCLLLLPSLAQIELLLHCLLLL
 LPSLAQIEVLLHCLLLLPSLAQIELLLHCLLLLPSLAQIEVLLHCLLLLAM*

Sequence 3279

Contig_0768_pos_1456_1800

is similar to (with p-value 4.0e-59)

>pir:pir|I67760|I67760 transposase (insertion sequence IS10
) - Escherichia coli >gp:gp|S67119|S67119_2 BST=somatotropin
 ...BST/beta-Gal fusion protein [Escherichia coli, LBB84, pla
 smid pXT107, IS10L/R-1, PlasmidTransposonInsertionMutant, 3
 30 genes, 1679 nt]. NID: g455674.
 atgcagattgaagaaaccttccgagacttgaaaagtccctgcctacggactaggcctacgc
 catagccgaacgagcagctcagagcggtttgatcatgctgctaatacgccctgatgctt
 caactaacatggttggttgcggcggttcacgctcagaaacaagggtgggacaagcacttc
 caggctaacacagctcagaaatcgaaacgtactctcaacagttcgcttaggcattggaagtt
 35 ttgcggcattctggctacacaataacaagggaagactcactcggtggctgcaaccctgctt
 actcaaaatctattcacacatggttacgttttggggaaattatga

Sequence 3280

MQIEETFRDLKSPAYGLGLRHSRTSSSERFDIMLLIALMLQLTCWLAVHAQKQGWDKHF
 40 QANTVRNRNVLSTVRLGMEVLRHSGYTITREDSLVAATLLTQNLFTHGYVLGKL*

Sequence 3281

Contig_0769_pos_8606_8947

is similar to (with p-value 6.0e-18)

>sp:sp|Q41364|SOT1_SPIOL 2-OXOGLUTARATE/MALATE TRANSLOCATOR
 45 PRECURSOR. >gp:gp|A47930|A47930_1 Sequence 1 from Patent WO
 9534654. NID: g2301793. >gp:gp|U13238|SOU13238_1 Spinacia ol
 eracea envelope membrane 2-oxoglutarate/malate translocator
 (SODiT1) mRNA, chloroplast mRNA encoding chloroplast protein
 50 , complete cds. NID: g595680.
 atgagtttgcttaaccatgggataaagcctaacttgtttaattgttctgccattaatata
 agaactgagaaccaaacgagtgattccatgcgcctgtttcatttaaaatatctgaccac
 gtaatacacctgtttaatagtaacaatgctaaagcaataaatgcagtgagcgtggcatca
 acattaatgaagcttcctaatacccaaaagccaatgctatgataaagacaccaaccatc
 55 aatttttcggctatagacatatgtcccatttcttctagttgttcagtagcccatTTTTTA
 gcgttaggcgtttctttaacagtaggtgggtataatttataa

Sequence 3282

MSLLNHGIKPNLFNCSAINTRTENQTSVFHAPVSFKISDHANTPVNSNNAKAINAVSVAS

TLMKLPNTHKANAMIKTPTINFSAIDICPISSSCSVAHFLALGVSLTVGGYNL*

Sequence 3283

Contig_0769_pos_9387_8242

5 >sp:sp|Q41364|SOT1_SPIOL 2-OXOGLUTARATE/MALATE TRANSLOCATOR
 PRECURSOR. >gp:gp|A47930|A47930_1 Sequence 1 from Patent WO
 9534654. NID: g2301793. >gp:gp|U13238|SOU13238_1 Spinacia ol
 eracea envelope membrane 2-oxoglutarate/malate translocator
 (SODiT1) mRNA, chloroplast mRNA encoding chloroplast protein
 10 , complete cds. NID: g595680.
 atggcctttttcatttcaagaggatttgtataaaacagggttaggtcgacgtattgctctg
 caattcgtaaattatttggaaagaaaacgcttggttggcttattcacttgttgggtgtt
 gaccttatcttagctcctgctacgccaagtaatacagcacgtgctggtggtattatgttt
 ccaatcattaagtccttgtcagagtcatattggttcatcgccgagagatggttctgagaga
 15 aaaatgggtgctgttttaattcttactgagttccaagtaatttaattacttcagctatg
 tttttaacagctatggccggtaaccctatagcgcaaagtttagctgaaaaaacggcacac
 gttcaaatcattacgatgaattggttgttgcgtgctattatacccggttgatttctctc
 atcggtgtcccttccattatttataaaattataccacactactgttaagaaacgcctaac
 gctaaaaaatgggctactgaacaactagaagaaatgggacatatgtctatagccgaaaaa
 20 ttgatgggttgggtgtctttatcatagcattggctttgtgggtattaggaagcttcattaat
 gttgatgccacgctcactgcatttattgcttttagcattgttactattaacaggtgtatta
 gcgtggtcagatattttaaatgaaacaggcgcatggaatacactcggttgggtctcagtt
 cttgtattaatggcagaacaattaaacaagtttaggctttatcccatgggtaagcaaactc
 attgctcaagggttgaaatggcttttagttggcctatcggttttagtttactcatctgttt
 25 tatttctactcacattatttatttcgcaagtgaacagcacatgtcagcgccatgtacgcc
 gcgttactcggtgttgcagtcgcttcgggtgcaccgccattattcagtgcattaatgtta
 ggggtcttttggttaacttactggcatcaacaacacactatagtagtgaccagcgccctata
 ttatacgcagatggctatgttacacaaaagcgctggtggactatgaatattgtacttgggt
 atagtctattttattatttggattggtgttaggttcactatggatgaaactcattggtatg
 30 atgtaa

Sequence 3284

MAFFISRGFVKTLGRRRIALQFVKLFGKKTGLAYSLVGVDLILAPATPSNTARAGGIMF
 PIIKSLSESPGSSPRDGSERKMGAFILIFTEFQGNLITSAMFLTAMAGNPIAQSLAEKTAH
 35 VQITWMNWFAAIIPGLISLIVVPFIYKLYPPTVKETPNAKKWATEQLEEMGHMSIAEK
 LVMGVFIIALALWVLGFSFINVDATLTAFIALALLLTGVLAWSDILNETGAWNTLVWFSV
 LVLMAEQLNKLGFIPWLSKLIAQGLNGFSWPILVLLILFYFYSHYLFASATAHVSAMYA
 ALLGVAVASGAPPLFSALMLGFFGNLLASTTHYSSGPAPILYADGYVTQKRWWTMNIVLG
 IVYFIIWIGVGSLSLWMLIGMM*

40

Sequence 3285

Contig_0771_pos_5868_0

is similar to (with p-value 1.0e-34)

45 >sp:sp|P45476|YHCC_ECOLI HYPOTHETICAL 34.6 KD PROTEIN IN AR
 CB-GLTB INTERGENIC REGION (F309). >gp:gp|U18997|ECOUW67_140
 Escherichia coli K-12 chromosomal region from 67.4 to 76.0 m
 inutes. NID: g606010.
 atgcatgaaaaatggagtgaaaggtcaatatattgcatactttcaggcggtttacaaatcgc
 catgcacctgttgaagtactaaaagaaaaatatgaacctgtcttaaaagaagatggcgctc
 50 gttgggtttatcaatcgcgcacaaagacctgattgtttgcctgatgatgtttagaataatta
 gctgaacttaatacagcgcaacttacttatgggtagaattggggcctacaaactgtgcacag
 tcaacttccgatttaataaatcggtgctcatgatatgcaaacatactatgacggcgtaaca
 aaattacgcaaacataataataatgtttgtacgcacataatcaacggcttaccaggtgaa
 aattatgacatgatgatggagactgctaaagaagtcgctcagatggacgttcaaggtatt
 55 aaaattcattttattacacttgctaaaaggaacgcctatgggttaacaatatgaaaaaggt
 atgctcagatttatg

Sequence 3286

MHEKWSEGQYIAYFQAFNTNTHAPVEVLKEYEVLKEDGVVGLSIATRPDCLPDDVVEYL

AELNQRITYLWVELGLQTVHQSTSDLINRAHDMQTYDGVTKLRKHNINVCTHIINGLPGE
 NYDMMMETAKEVAQMDVQGIKIHLHLKGTMPVKQYEKGMLEFM

Sequence 3287

5 Contig_0774_pos_1335_622

is similar to (with p-value 2.0e-89)

>sp:sp|P39149|UPP_BACSU URACIL PHOSPHORIBOSYLTRANSFERASE (E
 C 2.4.2.9) (UMP PYROPHOSPHORYLASE) (UPRTASE). >pir:pir|S4936
 4|S49364 uracil phosphoribosyltransferase - *Bacillus subtili*
 10 s >gp:gp|Z38002|BSSPORUPP_10 *B. subtilis* spoII-R, glyC and up
 p genes. NID: g556877. >gp:gp|Z99122|BSUB0019_186 *Bacillus s*
ubtilis complete genome (section 19 of 21): from 3597091 to
 3809700. NID: g2636029.

15 atgaagctaaagagagaggttcacgcactcacatctaaatatccattatataattaataga
 catacacattggaggaaaatgattatgagtaaagtacatgtttttgatcacccattaata
 caacacaaaactaagttatattagagatgctcgcactggaacaaaagagtttagggaactt
 gtagatgaagtcggtatgttaattggcttatgaagtaactagagacttagaactgcaagat
 gttgaaatacaaacacactgtgactaaaatgacagctaaacgttttggcgggtaaaaagtta
 gcaattgtacctattttaagagctgggtctaggtacagatggtgtgttaagtcttgtt
 20 cctgctgctagggtaggacatataggactatatagagatccagagactcttgaagcggt
 gactactttgcgaaaatgcctcaagacatcgatgaacgtcaaattattgtggtgatcct
 atgcttgcactactggtgcttcagctattgaagcaatttcttcattaaaaaacgtggagct
 aaaagtatacgttttatgtgttaatatagctgcccctgaaggcgttgaaaaaatgcaagaa
 gcacaccagatgtagatatatatattgctggcattagatgaaaaattaaatgacaaagcg
 25 tatattacaccaggtttaggtgatgcaggggatagattattcggtactaaataa

Sequence 3288

30 MKLKREFTHSHLNIHYI INRHTHWKRMIMSKVHVFDHPLIQHKLSYIRDARTGTKEFREL
 VDEVGMLMAYEVTRDLELQDVEIQTPVTKMTAKRLAGKKLAIVPILRAGLGMTDGVLSLV
 PAARVGHIGLYRDPETLEAVEYFAKMPQDIDERQIIVVDPMLATGASAIEAIISSLKKRGA
 KSIRFMCLIAAPEGVEKMQEAHPDVDIYIAALDEKLNDKAYITPGLGDAGDRLFGTK*

Sequence 3289

35 Contig_0775_pos_5036_4413

is similar to (with p-value 2.0e-43)

>sp:sp|P44310|KGUA_HAEIN GUANYLATE KINASE (EC 2.7.4.8) (GMP
 KINASE). >pir:pir|H64139|H64139 5'guanylate kinase (gmk) ho
 molog - *Haemophilus influenzae* (strain Rd KW20) >gp:gp|U3284
 8|U32848_2 *Haemophilus influenzae* Rd section 163 of 163 of t
 40 he complete genome. NID: g3212240.

atggataaggaaaaaggactgttaattgttctttcaggcccttcagggtgttgaaaggga
 actgttagaagaagatatttgaagacccaactacttcatataagtattctatatcaatg
 acgacacgtcatatgcgtgaagggtgaaattgatgggtgtagattacttctttaaacaag
 gaagaatttgaggcggttaattaaagacgaccagtttattgagtatgcacaatatgtaggt
 45 aattactatggtacacactgtacaatatgtaaaggatactatggaagaaggatcatgacgtc
 tttttagaatcgaagttgaagggtgctaagcaagtaagaaagaaatttccagatgcgttg
 ttcatattttttagcgcctccaagtttagatgacttgaaagaacgtcttgttggttagagga
 actgaatcagatgaaaagattcaaagtcgtgtgaacgaggcacgaaaagaagtagaaatg
 atgaatttatacgaactacgtttagttaaagcagcaggttgaaactcgctaagaatcgaatt
 50 cagtcattagttgaagctgagcatttaaaaagagagcgaatcgaagctaaatatagaaaa
 atgttactggaggtcaaaaaataa

Sequence 3290

55 MDKEKGLLIVLSGPGVGKGTVRKKIFEDPTTSYKYSISMTTRHMRERGEIDGVDYFFKTK
 EEFEALIKDDQFIEYAQYVGNYYGTPVQYVKDTMEEGHDFLEIEVEGAKQVRKKFPDAL
 FIFLAPPSLDDLKERLVGRGTESDEKIQSRVNARKEVEMMNLYDVVVVNDEVELAKNRI
 QSIVEAEHLKRERIEAKYRKM LLEVKK*

Sequence 3291

Contig_0779_pos_0_1180

is similar to (with p-value 6.0e-61)

>gp:gp|U31756|BSU31756_2 *Bacillus subtilis* gamma-aminobutyrate permease (gabP) gene and spore coat protein (cotA) gene, partial cds. NID: g969025.

5 atgattgccattgggggatgtataggaactgggtctattcatgacttctggaggagctata
catgacgcaggtgcattgggtgctttgattgcctatgcagttattggagcgatgggtgttc
tttctaatacgctcggttaggggagatggcgacatatattgcctgtgtcaggttcatttagt
acttatgtctacacgctttgtcgatccttcactaggttttgccttaggatggaattattgg
10 ttcaactgggtgataacagtagcagctgatgttactattgcagcgcaagttatacaatat
tgggtcccctatgcaaggtataccagcttgggtctggagttgtatttctctattattatt
ttcgcgcttaattctttatccgttagagtatatggagagagtgaatttgggttcgcactt
atcaaagtagttacagtcacatatttataggaattgggtatcttaactatttttagggatt
atgggtggagaattttaggatttgatacgtttacaaaaggagatgggccaataactaggt
15 gggaatttaggaggttagcttgcctatcaattcttgggtgtatttctagtcgcaggcttctca
ttccaaggaaactgaacttattggtattacagcaggtgaatctgaaaatccagaaagagca
gttccaaaagcgtattaaacaagtattttggcgtatacttttattttacattctagctatt
ttcattattgggtatggtgattccatagtagtaaggcattaatgggcgggtggtgatagt
atagctacttcaccttttacatttagtatttaagaatgctggattagcttttgcgtcttca
20 tttatgaatgctgttatattaacaagtgtattatcagcaggttaactcaggaatgtatgct
tcaacaagaatgttatattcgatgagtaaaagataaattagcttataattcttttggaaaa
acaaataaaaagtggtgctaccttatgtatctctaattgcaactggagtactagtcattctt
attttcgcattgcaacatttaagtggagatgcataatgaatacattgtagctgctagcgga
atgactgggtttattgcttgggttggtatagcaatcagtcacttttagatttagacgcgca
25 tttgataaacaatttatgataaatcaaaattaaaatatC

Sequence 3292

30 MIAIGGCIGTGLFMTSGGAIHDAGALGALIAVAVIGAMVFFLMTSLGEMATYLPVSGSFS
TYATRFVDPSSLGFGALGWNWYFNWVITVAADVTTAAQVIQYWSPMQGIPAWVWSCIFLII
FALNSLSVRVYGESEYWFALIKVVTVIIFIGIGILTILGIMGGEFVGFDFTFKDGPILG
GNLGGSLLSILGVFLVAGFSFQGTTELIGITAGESENPERAVPKAIKQVFWRILLFYILAI
FIIGMLIPYDSKALMGGDSIATSPFTLVFKNAGLAFASFMNAVILTSVLSAGNSGMYA
STRMLYSMSKDKLAYNSFGKTNKSGVPYVSLIATGVLVILIFALQHLSGDAYEYIVAASG
MTGFIAWVGIAISHFRFRRAFDKQNYDKSKLYX

Sequence 3293

Contig_0786_pos_2712_3827

is similar to (with p-value 7.0e-97)

40 >sp:sp|P96612|DDL_BACSU D-ALANINE--D-ALANINE LIGASE (EC 6.3
.2.4) (D-ALANYLALANINE SYNTHETASE). >gp:gp|AB001488|AB001488
_40 *Bacillus subtilis* genome sequence, 148 kb sequence of the
region between 35 and 47 degree. NID: g1881226. >gp:gp|Z99
106|BSUB0003_103 *Bacillus subtilis* complete genome (section
3 of 21): from 402751 to 611850. NID: g2632653.

45 atggaggaaaacgaaatgacaaaagaaaatatttgcctcgttttggaggtaaaagtgc
gaacacgatgtttcaattttaactgcacaaaatgttttaaacgcaattgataaagaacga
tatcaagttgatattcatttatatacaaacgatgggtgaatggaagaaaaagataatatt
acacaagaataaaaaatactgatgaactcgtcattaacgatgtagagactggagaaatc
tcacagttactcagtaaaaggttagtttaggaaaatcatatgatgcagttattccattattg
50 ctatggtccaaatggagaagatggaactatccaaggtcttttgaagtacttgatatacca
tatgtaggtaatgggtgtgttagctgcttcaagctcaatggataaaactcgtgatgaacaa
ttattttagcatagaggtttacctcaattaccttatatttagctttttaagaagtgagtat
gaaaaaatatgaaaataatcattaaattagttaatgataagttaacatatccggtattt
gtaaaacctgctaattctcggttcaagtggttggtataagtaaatgtaacaatgaagaagaa
55 ttaaaatctgggatagctgaagcattccaatttgatcgtaaaccttgctcattgaacaaggg
attaatgctagagagatcgaagtagctgtcttaggtaacgattatcctgaaacgacatgg
cctgggtgaggttggttaaggatgtagcgttttatgattataaatcaaggtataaagacggt
aagattagatttagattatccagcagatttagatcaagatgttcaaatgacattaagaaac
atggcatttagaggcctttaaagctactgattgttcgggatttagttcgtgcagatttcttt

gttactgatgataatcaaatttatattaatgaaacaaatgccatgccaggatttactgca
tatagtatgtatccaaacttatggaaaaatattgggcttatcttaccctgatttaattgct
aaattgattgatttagctaaagaacgttatgaagataaaaaagaaaaataatattcaciaa
atgtttcatctaaaaccgatcgtttttttatatttaa

5

Sequence 3294

MEENEMTKENICIVFGGKSAEHDVSILTAQNVLNAIDKERYQVDIIYITNDGEWKKKDNI
TQEIKNITDELVINDETGEISQLLSKGSLSYDAVFLLHGPNGEDGTIQGLFEVLDIP
YVGNGLAASSSSMDKLVMLKQLFEHRGLPQLPYISFLRSEYEKYENNIKLVNDKLTYPVF
VKPANLGSSVGISKCNNEELKSGIAEAFQFDRKLVIEQGINAREIEVAVLGNNDYPETTW
10 PGEVVKDVAFYDYKSKYKDGKIRLDIPADLDQDVQMTLRNMALEAFKATDCSGLVRADFF
VTDDNQIYINETNAMPGFTAYSMYPNLWKNMGLSYPDLIAKLIDLAKERYEDKKKNKYSQ
MFHLKPPIVFI*

15 Sequence 3295

Contig_0788_pos_6565_7542

is similar to (with p-value 8.0e-33)

>gp:gp|AB011003|AB011003_1 Candida albicans CaUAP1 gene for
UDP-N-acetylglucosamine pyrophosphorylase, complete cds. NI
20 D: g3273313.

atgaaaatttttgattatgaagatatacaattaattcccaataaatgtattggtgaaagc
agatctgagtgtaataacttcgggttaaatttgggcctcgtaacttttaaatggcagttggt
ccagcaaataatgcaaacagtcgatgaagaacttgacaaatggtttgcagaaaacgat
tatttttatatcatgcatagatttaataagaataatagaattccattataaaaaaatg
25 catcatgcagggttatttgccttctatttctggttgagttaaagaaaacgaatttaatttt
attgaaaaatttagcttcttcatcgtcataccagaatatataacaattgatattgctcat
ggctactcaaattcagttataaatatgattaagcatataaaaaaacatttaccaaattag
tttgtgatagctggtaattgttggagcgctgaaggagtaagagaacttgagaatgccggt
gcagatgctacaaaagtaggtattggtccagggaagagtattgattactaaaattaaaact
30 ggatttggtagcaggaggttggcaacttctgcgttaaattctttgtaataaggcagctaga
aaacctattattgcagatggaggttaagaacccacgggtgatataagccaaatcaattcgt
tttgggtgccactatggtaattggttcttatttgcgtgccacagaggaatcaccgggt
gaaaccgtcgagctagatggcaaaaaatataaagaatatttggtagtgctcagaatat
caaaaagggtgaacataagaacgttgaaggtaaaaaatgtttgtagaacacaaaggatct
35 cttaaagatacccttactgaaatggaacaagatttacagagttcaatttcatatgcagga
ggaaaggacttgaagtcattaagaacagttgattacgtcatagtaagaaattcaatcttt
aatggtgatagagattag

Sequence 3296

MKIFDYEDIQLIPNKCIVESRSECNSTSVKFGPRTFKLPVVPANMQTMNEELAQWFAEND
YFYIMHRFNEENRIPFIKKMHAGLFAISVGVKENEFNFIEKLASSSLIPEYITIDIAH
GHSNSVINMIKHKKHLPNSFVIAGNVGTPEGVRELENAGADATKVGIGPGRVCITIKT
GFGTGGWQLSALNLCNKAARKPIIADGGLRTHGDIKSIKIRFGATMVMIGSLFAAHEESPG
ETVELDGKKYKEYFGSASEYQGEHKNVEGKMFVEHKGSLKDTLTEMEDQLQSSISYAG
45 GKDLKSLRTVDYVIVRNSIFNGDRD*

Sequence 3297

Contig_0788_pos_4151_3402

is similar to (with p-value 6.0e-40)

>sp:sp|P27442|GUAC_ASCSU GMP REDUCTASE (EC 1.6.6.8) (GUANOS
50 INE 5'-MONOPHOSPHATE OXIDOREDUCTASE). >gp:gp|M82838|NEMGMP_1
Ascaris lumbricoides GMP reductase mRNA, complete cds. NID:
g159660.

atgacaagtgcatttaatacataaagatacaatagagtatttttaaacacataaatatttt
55 aactatgatgccaatcatattcatttctttaagcaagataacattggtgctttaagtga
gaaggaaaagcttgttttaataagagatggacataatggaacacctaattggttaagg
ggtgtattcaagtcctttaagaaagcaggataccttgataagatgcaacaagatcacgtc
aaatatactcttctaataacattgataatgtcttagttaagtttagaccggtattt
gccgggtttacagtgacacaaagtaagacatcacatcaaaaacaattcaacctaaagat

agtgaaagtgttaggtcggccttgtaaattgttgattgtaaagacactgtgttagagtattct
gaattagatagacagacatagcaaatcaatttaacaatgctaataataggtatccatgctttt
aaactaggtttcattaccagtgtgttgatagagaattaccgtatcatttggcaatcaag
caattaaagcaatttagatgaaaattttgggtgtggttgaaacgtcctacattaaagtttgaa
5 ttatttttattttgatataatttaggtacggtacaagctttgtaacattacaagttccaaga
gaagaagaattttcaccttttaaaaaataaagaaggttaaggataggtgtacatacagctaca
gaagattttaaacgaatggatttgatttaa

Sequence 3298

10 MTSDINHKTIEYFKQHKYFNYDANHIHFFKQDNIVALSEEGKLVNLRDGHIMETPNGNG
GVFKSLKKAGYLDKMQQDHVKYIFLNNIDNVLVKVLDPLFAGFTVTQSKDITSKTIQPKD
SESVGRLVNVDCKDTVLEYSELDTDIANQFNNANIGIHAFKLGFIITSVDRELPHYHLAIK
QLKQLDENFGVVERPTLKFELFYFDIFRYGTSFVTLQVPREEEFSPLKNKEGKDSVHTAT
EDLKRMDLI*

Sequence 3299

Contig_0795_pos_2866_4128

>gp:gp|AJ224946|CGA224946_1 Corynebacterium glutamicum DNA
for L-Malate:quinone oxidoreductase. NID: g3059092.

20 atgcacatgagtgaaagcaaatcataaaacatcggtgttgtaggtgcaggaattatttggt
acgtcagtagcgacaatgctttcaaaagtaagtcctaactggcatatcgatatgtttgaa
agactagaaggcgctggtattgaaagttcaaatgaaaataataatgctgggacaggtcat
gcggcattatgtgaattaaactatacagttgaacaagatgatggttcaattgatgcatct
aaagcgcaagaaattaatgaacaattcgaattatctagacaattctggggtaatttagtt
25 aaaaatggtgatatttctaattcctgaagaatttattcaaccattacctcatatcagtttc
ggtatgggaccaacaaacgttaactttttaagaaaacgttatgaaacactaagaactctt
ccaatgttcgatacaatcgaatatacagaagacatggaaacaatgagaaaatggatgccat
ttaatgatggaaaatcgtgaaccaggtcatcaaatggcagcaagtaaaattgatgaaggt
acagatgtgaactatggtgcgttaacacgtaagtttagcacattacttagaacaataatct
30 aatgtttcattaaaaatacaatcatgatgtttgtagatttaacacaaagagaagatggcaaa
tgggaagttgtcgttgaaaatagagaaaactaaagaaaagtaactaaaatagcagataaaa
gtgtttatttggtgctggcggtcactctattccgttattacaaaaatctggcggttaacaa
agagaacaccttaggtggtttcccaatcagtggtcaattcttaagatgtacaaaccagat
attattaaacaacatgcggctaaagtttacagtaaaagagcctcaaggttaagccaccaatg
35 actgtaccacaccttgatacaggttatatcaatggtaaaacaacattattattttggtcca
tatgcgaatatcgggccctaaattcttgaaattcgggttcaaatctagacttattcgaatca
atcaaacatataacattactacaatgttggttcagcagttaaaaatgtacctttaatt
aatatttcaattgatcaaatgataaaaactaaagaaggttgatgaactatttaagaaca
tttattctgatgctaaagatgaagattgggaactttacactgctggttaacgtgttcaa
40 gttattaaagataaccatttgaagcaggaccttactttattgcagcagtcgtatttgta
tag

Sequence 3300

45 MHMSEANHKNIIVVVGAGIIGTSVATMLSKVSPNWHIDMFERLEGAGIESSNENNNAGTGH
AALCELNYTVEQDDGSIDASKAQEINEQFELSRQFWGNLVKNGDISNPEEFIQPLPHISF
VMGPTNVNFLRKRYETLRTLPMFDTIEYTEDMETMRKWMPLMMENREPGHQMAASKIDEG
TDVNYGALTRKLAHYLEQKSNVSLKYNHVDVLDLQREDGKWEVVVENRETKEKVTKIADK
VFIGAGGHSIPLLQKSGVKQREHLGGFPISGQFLRCTNPDIKQHAQVYSKEPQKPPM
TVPHLDTRYINGKQTLFGPYANIGPKFLKFGSNLDFESIKPYNITMLASAVKNVPLI
50 KYSIDQMIKTKEGCMNYLRTFIPDAKDEDWELYTAGKRVQVIKDNHLKQDLTLLQQSYLL
*

Sequence 3301

Contig_0795_pos_4741_5160

55 is similar to (with p-value 7.0e-17)

>sp:sp|P77279|YBBL_ECOLI HYPOTHETICAL ABC TRANSPORTER ATP-B
INDING PROTEIN IN USHA-TESA INTERGENIC REGION. >gp:gp|U82664
|ECU82664_88 Escherichia coli minutes 9 to 11 genomic sequen
ce. NID: g1773084. >gp:gp|AE000155|AE000155_6 Escherichia co

li K-12 MG1655 section 45 of 400 of the complete genome. NID
: g1786692.

5 atgcaacaaagtgagttaatcgggtataacaattgaagataatatgaaatttctctgctgag
gctagaagtgaaagcttttgaccgtgataaagcgaaacaactcatctctcaagtaggatta
ggtaattatcagtttagatgctcaaattgagcacatgtctgggggagagcaacaacgtatt
accatcgctagacaactcatgtatgaacctgaagttttattattggacgaagctactagc
gcttttagatacacataataaaaaagaaaattgaagaaatttatatttaaactagcagataaa
gggattgccattttgtggattacgcatagtgatgaccaaagtatgcgtcattttaagcgt
agaatcacaattactgacggttaagatatcgagtgatgaggagttgaatggtaatgagtaa

10 Sequence 3302
MQQSELIGYTIEDNMKFPAEARSEAFDRDKAKQLISQVGLGNYQLDAQIEHMSGGEQQRI
TIARQLMYEPEVLLLEATSALDTHNKKKIEEIIIFKLADKGIAILWITHSDDQSMRHFRR
RITITDGKISSDEELNGNE*

15 Sequence 3303
Contig_0795_pos_5525_5929
is similar to (with p-value 2.0e-32)

20 >sp:sp|P77307|YBBM_ECOLI HYPOTHETICAL 28.2 KD PROTEIN IN US
HA-TESA INTERGENIC REGION.
atggttagctaataatgggttaatacgcgattaatctcgcttatcagaatttagaaagagca
tttggttaagatgtttctgatattgaatccaaacttacgttagcagcgacacctaagctc
gcataaaaatcagctattagagaaagtatacgcttagcaattgttcctacaattgattct
gtaaaaacatatggctagtttcaattccaggtatgatgacaggattgattatcggaggc
25 gttgacccacttcaagcaattaaatttcaattgcttgctgctgtttattcatacaacagcg
acgattatgtctgcactcattgcaacgtatatgagttacgggtcagttctttaatgctcgt
catcaactcattgctagaacgcaacgcacacaagacaagtagttaa

Sequence 3304
30 MLANGLIAINLAYQNLERAFVQDVSDIESKLTLAATPKLASKSAIRESIRLAIVPTIDS
VKTYGLVSIIPGMMTGLIIGGVDPLQAIKFQLLVVFIHTTATIMSALIATYMSYQFFNAR
HQLIARTQRTRQSS*

Sequence 3305
35 Contig_0795_pos_6068_7246
is similar to (with p-value 7.0e-24)
>gp:gp|A08113|A08113_1 Synthetic DNA sequence of chloramphenicol-resistance gene. NID: g413362.

40 atgactgtcgcacgcacaacaacatttatattaagtgttttcacgttaggtatgggtgag
atgatgggtgcaggattatgaatttaagtgaatcattcatgtttctgaagcgggtg
gttgggcaactggtaacgttgtagcgcgttacctttgcgatatgtggtccgttattagt
aagttgactcatcggtttacttcgcgatcagttattatggacgttaattgtctttatc
ttcgggaatgggtatgattgccattgcacctcattttggaataatagttgtaggacgtatt
45 ttatcttctgcgcagcttcactcattattgtgaaagttttagcactcacagcgatgctc
acatcagcaaaaaacagaggtaaaatgattgggtattgtttatacaggttttagtggggcg
aatgtctttgggtgttcccatcggtacagtgattggcgactgggtaggatggcgatttaca
tttttattcattattattgtaagtgtatttgttggtgttttaattgtaatttatctacca
aaagaagatgaattgtcacatccaaatcaaacacctcggtcatctagtattgaatcaca
actggctcaagcgtcataagacctcggtgaggtttttaaatatctgatgattacattttta
50 gtgctgggtgctaattctgtaacattcgtgtttattaatccatttaattttatccaatgga
catgaaatgtctttgtgtcttttagcactactgttaattgggtgtagcagggtggtgattgg
acttcattagggcgggtgttttatctgataagtttactagtaagcgttggttaataatttcg
atttcaatattttataataatgatgataattcttaacttattattaccaggaacaggatta
ttattagttggctattttatgtggaatttaattgcagtgaggtacaaatccagctattcaa
55 agtgggtattattgaacacgtcgaaggagatacaagccaagtgatgagttggaatatgtcg
agtctcaatgccggtattggcgttggcggaatcgtagggtgactcgttatgacacattta
tcagtggaaatggttacttatactagtgacattgataggtttaattagttcttatcattgta
ttcacttttaaaaaatagacattatgctaaaaatttatga

Sequence 3306

MTVARTTTFILSVFIVGMVEMMVAGIMNLMSQDLHVSEAVVGQLVTLYALTFAICGPLL
KLTHRFTSRVLLWTLIVFIFGNGMIAIAPHFGIIVVGRILSSAAASLIIVKVLALTAML
TSAKNRGMIGIVYTGFSGANVFGVPIGTVIGDWGWRFTFLFIIIVSVFVGVLMMLIYLP
5 KEDELSHPNQTPRSSIESQTGSSVIRPREVFKYLMITFLVLVANSVTFVFINPLILSNG
HEMSFVSLALLVNGVAGVIGTSLGGVLSDKFTSKRWLIISISIFIIMMIILNLLPGTGL
LLVGLFMWNLMQWSTNPAIQSGIEHVEGDTSQVMSWNMSSLNAGIGVGGIVGGLVMTHL
SVEYVTTYTSALIGLISLIIVFTLKNRHYAKNL*

10 Sequence 3307

Contig_0799_pos_3893_3078

is similar to (with p-value 1.0e-20)

>gp:gp|U67964|EVU67964_2 Ectromelia virus H14-B and H14-E g
enes, complete cds. NID: g2145123.

15 atgcattatataaaattttattgagtcaaaagataatacaaaacttttatgaaagtgaat
gatattcaagatgcaaaagcgaatatcattatagctcatgggtgtggcagaacatttagat
cgttatgatgagataacagcatatttaaataagcgggttttagtggtatttagatatgat
caaagaggcgatgggtcggtctgaaggcaagcgtgccttttatagcaatagtaataaatt
gtcgaagatttagatgcatgataataaattatgtgaagtcaaaactttgaaggtaaaagttac
20 ttaatcggctcatagtatgggtggttatacagtcacttttatatggaacgaaacatccaaat
acagtgaatgggtattataacttctggagcattaacacgttataataataaactatttggc
aatcctgatagaaacatatcacctgatacttatatagaaaacaatttaagtggggggtta
tggtctgatttagaggaatggaaaaatataaacttgatgatttgaaatgcaaaacaatc
tctatggggctcgtcttttcaataatggatgggtggttaggtatttgaaagacaatgctcaa
25 caatttacagataaatattttgatattgcattggcaaggaagatgggctagtaagctatgta
gattctttacagatctttatcaagaaataggatcagcacataaatcattacacatctatgat
cgtttgagcatgaaatatttaataagagttcttataatagaactatttttaacgaagtt
attgaatggcttgaaacggaattaacttataactaa

30 Sequence 3308

MHYIKFIESKDNTKLYMKVNDIQDAKANIIIAHGVAEHLDRYDEITAYLNEAGFSVIRYD
QRGHRSEGKRAFYSNSNEIVEDLDAIINYVKSNEFGKVYLIGHSMGGYTVTLYGTKHPN
TVNGIITSGALTRYNNKLFGNPDNRNISPDTYIENNLSEGVCSLEVMKEYKLDDLNAKQI
SMGLVFSIMDGVRYLKDNAQQFTDNILILHGKEDGLVSYVDSLQLYQEIGSAHKSLHIYD
35 RLEHEIFNESSYNRTIFNEVIEWLETELTYN*

Sequence 3309

Contig_0802_pos_4926_3982

is similar to (with p-value 3.0e-73)

40 >gp:gp|S72926|S72926_1 Hordeum vulgare glucose and ribitol
dehydrogenase homolog mRNA, complete cds. NID: g633889.
gtgtgggcttggtacactcattacaagaattcaattctgatccgaatatacaaaatttt
gatgaaatgcttaacaaactacaaaagggtcagtttaatttagtgcaagtgaagaaagtggtg
actaaaaaaattgtagatcactttgtcgaagaattatagcgaagaacaaaacaaaaa
45 atcaatacaggttataaaactggtggattacaaaataggtggtttagaacctacacagttg
attgtaatcgctgcgagaccgtcagtaggtaaaaacggggtttgcgcttaatatgatgctt
aatatagcgtctcaaggctataaaaacttcattcttcagtctagagacaactggcgtgtct
gtattgaaaaggatgttatcagcagaaactgggtagaactaactcgtatcaaagaaatt
aaagatttagaacgggatgaattaacacgtttaacaactgcagcagacagaataactcaaa
50 cttgatatagatatacacgataaaaagcaatattactacacatgatgtacgtaaaacaagcg
atgaagaacaaagatgtgcaacaggttatcttcattgactacttacaacttatgcagaca
gacagtaagttagatcgctcgtaattggtatcgaaaagatcgcgagatttgaagattatt
gcaaatgaaacaggtgcaattattgtgttgctatctcaattgagcagaggtgtagaaaca
agaaatgacaaaagacctatgctatctgacatgaaagaagcaggtggaattgaagcagat
55 gcaagtttagctatggtgttatatcgagatgattactacaaccgtgatgatgttgatgac
tcaggcaagtcattgttgatgtaacatcgcaagaataaagacggagaaacaggtgta
gttgagtttgagtactacaagaaaacgcagaggttcttcacatga

Sequence 3310

VWAWYTHLQEFNSDPNIQNFDEMLNKLQKVSLISASEESGTTKIVDHFVEELYSEEPKQK
 INTGYKLVYKIGGLEPTQLIVIAARPSVGKTGFALNMLNIASQGYKTSFFSLETTGVS
 VLKRMLSAETGIELTRIKEIKDLEPDELTRLTTAADRIKLKLDIDIHDKSNITTHDVRKQA
 MKNKDVQQVIFIDYLQMLQTD SKLDRRNGIEKISRDLKIIANETGAIIVLLSQLSRGVET
 5 RNDKRPMLSDMKEAGGIEADASLAMLRYRDDYNNRDDVDDSGKSIVECNIAKNKDGETGV
 VEFYYKKTKQRFFT*

Sequence 3311

Contig_0802_pos_3039_2170

10 is similar to (with p-value 3.0e-36)
 >sp:sp|P37469|DNAC_BACSU REPLICATIVE DNA HELICASE (EC 3.6.1
 .-). >gp:gp|D26185|BAC180K_4 B. subtilis DNA, 180 kilobase r
 egion of replication origin. NID: g467326. >gp:gp|Z99124|BSU
 B0021_149 Bacillus subtilis complete genome (section 21 of 2
 15 1): from 3999281 to 4214814. NID: g2636442.
 gtgaattttataaaatttcacaacaaaatcaaaggatatactcaaaatagacaaccaggt
 attgaagcggatattggaacctaaacccattgcagaattagaagaatataaagcagcagga
 aagttagagaataaagtgtgctctaataacaggaggagattcaggtattggacgtgcgata
 gcaatactatatgtctaagaaggggcacaatgttgctattggttattatgacgaacatcaa
 20 gatgccgaagacacagtttaatcgacttcaagaaatgggtgtaaaagctaaagcttatgct
 catgactttaaagatgaaaagcaatctcaaaagttatcaaagatgtcataaatgacttc
 ggtagtttaaatatattagtaaaatggtggcgctgcaatttccacgcgatcattttgaa
 gatatactccacaacaagtgaagaaaacttttatgacgaatatttttggtatgatgttt
 ttatcccaatcagcagtagcttacctatctgaaggagatacaattataaaactacaagt
 25 gtcacagcatatagaggatcggggcattctcattgattattcagctacaaaagggtgcata
 gtatcggtttaccggttctcttgctactactttaatggaaaagggaattcgcgttaacgcc
 gttgccccggcccaatctattcacctttaattcctgcgacatttgatgaagaaaaagta
 gaacatcaagggtgacgaaacgccgatgggtcgctcgtagacaaccagcagaacttgacact
 tcttatgtctcttagcaacacatgcagatagttcctatattactgggtcaagtcattcat
 30 gtcaatggtggcgatttttatcacatcttaa

Sequence 3312

VNLLKFHNKIKGYTQNRQPGIEADMEPKPIAELEEYKAAGKLENKVALITGGDSGIGRAI
 AILYAKEGANVAIGYYDEHQDAEDTVNRLQEMGVKAKAYAHDLKDEKQSQKLIKDVINDF
 35 GSLNILVNNGGVQFPRDFEDITPQQVKETFMNTNIFGMMFLSQSAVPYLSGDTIINTTS
 VTAYRSGHLIDYSATKGAIVSFTRSLATTLMEKGIRVNAVAPGPIYSPLIPATFDEEKV
 EHQQDETTPMGRRGQPAELAPSYVFLATHADSSYITGQVIHVNGGDFITS*

Sequence 3313

40 Contig_0804_pos_4400_3708
 is similar to (with p-value 3.0e-62)
 >sp:sp|P96612|DDL_BACSU D-ALANINE--D-ALANINE LIGASE (EC 6.3
 .2.4) (D-ALANYLALANINE SYNTHETASE). >gp:gp|AB001488|AB001488
 _40 Bacillus subtilis genome sequence, 148 kb sequence of th
 45 e region between 35 and 47 degree. NID: g1881226. >gp:gp|Z99
 106|BSUB0003_103 Bacillus subtilis complete genome (section
 3 of 21): from 402751 to 611850. NID: g2632653.
 atggataaactcgtgatgaacaattatttgagcatagagggttacctcaattaccttat
 attagctttttaagaagtgtgatgaaaaatatagaaaataatattcattaaattagttat
 50 gataagttaacatatccggtatttggtaaaacctgctaactcgggttcaagtgttggtata
 agtaaatgtaacaatgaagaagaattaaaatctgggatagctgaagcattccaatttgat
 cgtaaaactgttcattgaacaaggattaatgctagagagatcgaagtagctgtcttaggt
 aacgattatcctgaaacgacatggcctgggtgaggtgttaaggatgtagcgttttatgat
 tataaatcaaagtataaagacggtaagattagattagatattccagcagatttagatcaa
 55 gatgttcaaatgacattaagaacatggcattagaggcctttaagctactgattgttcg
 ggattagttcgtgcagatttcttcttactgatgataatcaaatttatattaatgaaaca
 aatgccatgccaggatttactgcatatagtagtatccaaacttatggaaaaatatgggc
 ttatcttaccctgatttaattgctaaattgattgatttagctaaagaacgttatgaagat
 aaaaagaaaaataaatataaaattgattattag

Sequence 3314

MDKLVMKQLFEHRGLPQLPYISFLRSEYEKYENNIKLVNDKLTYPVFVKPANLGSSVGI
SKCNNEEELKSGIAEAFQFDRKLVIEQGINAREIEVAVLGNDYPETTWPGEVVKDVAFYD
5 YKSKYKDGKIRLDIPADLDQDVQMTLRNMALEAFKATDCSGLVRADFFVTDNQIYINET
NAMPGFTAYSMYPNLWKNMGLSYPDLIAKLIDLAKERYEDKKKNKYKIDY*

Sequence 3315

Contig_0808_pos_6745_6269

10 is similar to (with p-value 1.0e-55)

>gp:gp|AF068246|AF068246_1 Mus musculus SA protein mRNA, complete cds. NID: g3928675.

gtgtttaattacttctacgaaagaagacgaaagagaatgggttgaagcaattagagta
gcaagaaatatcctaaacaaaaagctatggaccatttaattggtggcgaaatttcacca
15 ggaccacaagttcaaacggatgaagaaattctagattgggtacgtaaagatggagaaact
gcattacatccatctttagcgcgaaaaatgggacctgcattcgtacccaatggcagtagtc
gatccattaactatgaaagtacatggtatggaaaatttacgtgctggtgatgcttcagca
atgcctagaacaacaaatggtaattattcatgcacctgtattgatgttagctgagaaagca
gcggacattattcgtggtagaaaaccgcttgaacctcaatatgttgactattataaacat
20 ggtattgatgatgaaaaagcaggtgcaatggaagatgatccattctaccaatattaa

Sequence 3316

VFNYLSTKEDEREWVEAIRVARNILKQKAMDPFNGGEISPGPQVQTDEEILDWVRKDGET
ALHPSCSAKMGPPASDPMAVVDPLTMKVHGMENLRVVDASAMPRTTNGNIHAPVLMMLAEKA
25 ADIIRGRKPLEPQYVDYYKHGIDDEKAGAMEDDPFYQY*

Sequence 3317

Contig_0808_pos_3296_2034

30 is similar to (with p-value 7.0e-29)

>sp:sp|P17444|BETA_ECOLI CHOLINE DEHYDROGENASE (EC 1.1.99.1)
(CHD). >gp:gp|X52905|ECBET_5 Escherichia coli betT, betI,
betB and betA genes. NID: g48714. >gp:gp|M77738|ECOBETA_1 E.
coli choline dehydrogenase (betA) gene, complete cds. NID: g
145401. >gp:gp|AE000138|AE000138_2 Escherichia coli K-12 MG1
35 655 section 28 of 400 of the complete genome. NID: g1786501.

atgaataaatcaaatctactagcacctgagaattataatattgttacagaaatagaaaaa
tatgcctcagaagatcataaaaaagccattatttacaaggataacgagcatgaaaaatatt
tctgtaagttataaagaacttatcagtaatgctaataaagtagggaatgtattcctcaat
40 catgggctaaaaaaggagataaagttctcatcatgatgccacgtgcaatcggttacatat
gaattatatattgcagcattgaaactagggatagcgattgttccaagttcggaatgtta
cgaacaaaagatttacaatatcgaattactcacgggtgagattgatgcagttatttcattt
gattctctaactaaagaatttgaacggttaagaatattgaccaattaaaaaaatttata
gtagctggtcacaaagaagattgggtttcaatagaagatgaaaaagaaaaagtaagtgat
45 gaccttaaggcgagatacaacacgagatgatttggcgattcttcttatacatcaggt
acaacaggcaatccaaaagcagtaacgcattcacatggatgggggtatgccatttaciaa
atggcaccacaaacattggttatgtatacaagagaatgatcttgtatgggcaactgcagca
ccagggtggcaaaagtggtgtggagtcatttttatctgtattagggatgggagcaaca
gcatttgtctataacggtcgtttccaccctgaaacatatctcgagttacttcaaaattat
50 caaattaatgttctatgttgtacaccaacagaatatcgatgatggctaaacttagtcat
ttagaacagtacaatttagagtatttaccacagtgcggtgtctgcgggtgaacctttaaat
cgagaagtgttgaacaattttaaaccgtcattttaaattactgttcgagatggatatgga
caaaccgaaagtacattgttgatcggtatttctaaaagatactgaaccacgtatgggttct
atgggcaaaaggtatacctggttagtttcttactgtcattgacgatgatggtaaagaggtt
55 ggtccaaatgttaaaggtaatatcgccgtgcctttagacttaccggctttattttaaagggt
tactttaaagatgaagcacgcacaaaagcagcttcaacaggtgattattatgttactgga
gaccaagctcatattgatatatatcagtggtcaatcatttttacgtcctttattttaaaaa
taa

Sequence 3318

MNKSNLLAPENYNIVTEIEKYASEDHKKAI IYKDNEHENISVSYKELISNANKVGNVFLN
 HGLKKGDKVLIMMPRAIVTYELYIAALKLGLAIVPSSEMLR TKDLQYRITHGEIDAVISF
 DSLTKEFENVKEYDQLKKFIVAGHKEDWVSI EDEKEKVSDDLKGADTTRDDLAILSYTSG
 5 TTGNPKAVTHSHGWGYAHLQMAPKHWLCIQENDLVWATAAPGWQKWVWSPFLSVLGMGAT
 AFVYNGRFHPETYLELLQNYQINVLCCTPTEYRMMAKLSHLEQYNLEYLHSAVSAGEPLN
 REVVEQFKRHFNITVRDYGQTESTLLIGFLKDEPRMGSMGKIPGSFVTVIDDDGKEV
 GPNVKGNI AVPLDLPALFKGYFKDEARTKAASTGDYVVTGDQAHIDIYQWSIIFTSFILK
 *

10

Sequence 3319

Contig_0810_pos_827_1459

is similar to (with p-value 8.0e-31)

>gp:gp|D64024|D64024_2 Sulfolobus sp. DNA for 2-oxoacid:fer
 15 redoxin oxidoreductase subunit alpha and beta, complete cds.
 NID: g1565182.

atggcaaaacaaagatttaacagttatcgcttctggtggtgatggagacggctatgcaata
 ggaatgggacatactattcatgctcttagacgtaatatgaatatgacgtatattgtcatg
 gacaatcaaataataggattaactaaaggacaaacatcaccttcctcagctaaaggattt
 20 gtaactaaatcaaacctaaaggaaatatagaaaaagaatgtagctccattggaattggca
 ctgtcctctggtgcaactttttagcacaaaggattctcaagtgatataaaggcattaact
 aaaatgattgaagatgcgattcatcatgatggtttttcttttgtaattgttttctcacct
 tgtgttacttacaataaagtgaatacttatgactggtttaagaacatttaacaagtatc
 gatgatattgagggctatgacatcacagataaacaacttgctatgaaaactgtgctggat
 25 catgagtcactggttaaaggtatcgtttatcaagatacaacaacaccttcttatgaatcg
 caaatttcagaactagaacatgagggcgttagctaaaagagatattcatattacagaagaa
 actttcaacgatttaactgcacaatttttataa

Sequence 3320

MANKDLTVIASGGDGDGYAIGMGHTIHALRRNMNMTYIVMDNQIYGLTKGQTSPPSSAKGF
 30 VTKSTPKGNIEKNVAPLELALSSGATFVAQGFSSDIKALTKMIEDAIHHDGFSFVNVSFSP
 CVTYNKVNTYDWFKEHLTSIDDIEGYDITDKQLAMKTVLDHESLVKGIVYQDTTTPSYES
 QISELEHEALAKRDIHITEETFNDLTAQFL*

35 Sequence 3321

Contig_0812_pos_620_1003

is similar to (with p-value 3.0e-63)

>pir:pir|A43577|A43577 regulatory protein pfor - Clostridiu
 m perfringens

atgactgctaataatgcatatgggttggttatcacctactgaagtgattgcactaccaatttgt
 40 ttaagtactgcaataacccttgagaaattaagtttgcaagaccaaataactgctgga
 gccataaaaataacaacaattaaatctagtcctctggaactttttctcaacaaacttg
 attccaaatgcaacaacatagctgcgataaacgcaggtaataattttgagtcattgta
 actaaacctacaatgactgcaataactggagacacttctaacttttagacaagttaata
 45 ccaacagctataccacttaaaactccctgctagatcaccaatttcttgaaaaatttaaca
 tggataacgccaccaattgcgtaa

Sequence 3322

MTANAYGLLSPTTEVIALPICLSTAITPGEIKFARNVNTAGANKITTIKSSPSGTFSTNL
 50 IPNATTYAAINAGNNFESCNTKPTMTANTGDTSNFRQVNIPTAIPKLPARSPISWKNLT
 WNTPIIA*

Sequence 3323

Contig_0812_pos_2366_5200

is similar to (with p-value 2.0e-07)

>sp:sp|O34863|UVRA_BACSU EXCINUCLEASE ABC SUBUNIT A. >gp:gp
 |Z99122|BSUB0019_13 Bacillus subtilis complete genome (secti
 on 19 of 21): from 3597091 to 3809700. NID: g2636029. >gp:gp
 |AF017113|AF017113_13 Bacillus subtilis 300-304 degree genom

ic sequence. NID: g2618830.

atgaaaggaccgtcaattgtagtaaaagggtgcacgagcacataattttaaaggagtagat
 attgaattacctaataaacaagttgattgtcatgactggactttcaggctcaggtaaatct
 tcgttagcgttcgatactatttatgctgaagggtcaacgacgttatgtggaatcggttaagt
 5 gcatatgcgcgacaaattttaggacaaatggacaaacctgatgtagataccattgaagggt
 ttgtcaccagcaatatctattgatcagaagacgacaagtaaaacccacgttctaccggt
 gccactgtgactgaaatctatgattatattcgattattatgctgcgtgtcggtaaaccg
 tactgtccttatcacggtatagaaattgaatcacaaactgttcaacaaatgggtgatcgt
 attttagaattggaagaacgtacaaaaattcaattacttgcctcagttatatcacataga
 10 aaaggaagtcatagaaatgaattgaagatataggtgaagaaaggatatgtgcgtttacgt
 gttgatgatgaaatcgctcgatgtaaatgaagtgccctcaattggataaaaaataaaatcac
 acaattgaagttggttgtagtagactttagttaaagatggcattgagacgcgttttagcg
 gactctattgagacagcacttgaacttgcgtgaaggtaatttaacagttgacgttattaat
 ggagaagaacttaaaattctctgagaaccacgcatgtcctatatgtggtttctctattggt
 15 gaattagaacctagaatgttttagtttcaacagtcggtttgggtgcgtgtccaacttgtgat
 ggttttaggtcaaaaactaaaagtagacctagatttagtcatctctgataaaaaataaaact
 ttaaatgaaggtgcaattgaaccatgggaaccaacaagttcggatttttatccaacttta
 ttaaacgctgatatgtgaagtcacaaaataaatatggataaagccatataaaaaattgact
 gatagacaaaaaataatacttatgaatgggttctggagaaaaagaaattgaatttactttc
 20 acgcaacgaaatgggtggtactcgtaaacgtaaaatgggttttgaagggtggttacctaac
 attgaccgtcgttaccatgaatctcctctgaatatacacgcgaaatgatgagtaaatat
 atgacagagttaccgtgtgaaacttgcctatggcaaacgcttaagcaagaagcgttatct
 gtatacgttaggcgactataatataaggtgaagtcgttgaatattcaattaaaaacgcgctt
 tattatttcgaaaatttaaaattaaagcgaccaagataaatcgattgcagatcaaatttta
 25 aaagagattatttcaagattatcatttttaataatgttgggttggaaatttgacttta
 gatcgctcatcagggaactttatctggagggaagcacaacgtattcgttttagcaacacaa
 attgggttcacgattaaccggagtagctatatgttttagatgaaccttctattggattacat
 caaagagataatgatagattaatttaatacttttaaaagaaatgcgtgatttaggtaataca
 ctatttgtcgttgaaatgacgatgatactatgagagcagctgattatttagttgatgtg
 30 ggtccgggagctggttaaccacggtggagaggttgcctcaagtggtagccctaataaagta
 atgaaagataaaaaatccttaactggtcaatatttaagtggaaaaaacgaattgaagtc
 cctgaatacacgacgagaaatcacgtagaagattcaaatataaggtgctaaaagtaaat
 aatttgaaaaatgtaaatgtagacttcccactatctgtcttaactgttggtacaggtgtg
 tcagggttctggtaaaagttcactcgtcaatgaaattttatataaagcatttagctcaaaaa
 35 attataaatactaaagtgaacacctgggaattttgatgaaattaaaggaattgatcaatta
 gataaaatcattgatattgatcaatcgccaataggtagaacaccacgttctaaccagcc
 acatacactgggtgtctttgatgacataagagatgtctttgcacaaacgaatgaagctaaa
 atacgaggttatcaaaaaggtagatttagtttaattgtcaaaggtggacgatgtgaagct
 tgtaaaggtgatggaattataaaaaattgaaatgcattttttaccagatgtctatgtacct
 40 tgtgaagtatgtgatggtaaacgctataatcgtgagacttttagaggtaacatacaaaggt
 aaaaatattgcggatgtattagaatgactgttgaaagagctacgcatttctttgaaaat
 attcctaagattaaacgtaaatatacaaacacttgtatagttgggttgggttacattact
 ttaggtcaacaagctactacattatctggtggcgaagcgcaacgtgtaaaactcgcacatca
 gaattgcacaaacgttcaacggggcgttctatttatattcttgatgaaccaactacagga
 45 ttacatgtcgcagatataagtcgtttattaaaggtattgaatcgtagtggaatgggt
 gatagcgtcgttattatcgaacacaatcttgatgttattaaaacggctgatcatattatt
 gatttaggtccagaaggcgtgaaggtggaggaacaatcatcgcaactggtacacctgaa
 gagattgctcaaaataaagggtcttacactgggtcaatacttaaaaccagtttagagaga
 gacagcgttgaatag

50

Sequence 3324

MKGPSIVVKGARAHNLKGVDIELPKNKLIIVMTGLSGSGKSSLAFDTIYAEGQRRYVESLS
 AYARQFLGQMDKPDVDTIEGLSPAISIDQKTTSKNPRSTVATVTEIYDIYIRLLYARVGKP
 YCPYHGIEIESQTVQQMVDRIELEERTKIQLLAPVISHRKSHEKLI EDIGKKGYVRLR
 55 VDDEIVDVNEVPQLDKNKNHTIEVVVDRLVVKDGIETRLADSIETALELAEGNLTVDVIN
 GEELKFSENHACPICGFSIGELEPRMFSFNSPFGACPTCDGLGQKLKVDLDLVI PDKNKT
 LNEGAIEPWEPTSSDFYPTLLKRVCEVYKINMDKPYKKLTDRQKNILMNGSGEKEIEFTF
 TQRNGGTRKRKMVFEGVVPNIDRRYHESPSEYTREMMSKYMTLPCETCHGKRLSKEALS
 VYVG DYNIGEVVEYSIKNALYFENLKLSDQDKSIADQILKEIISRLSFLNNVGLLEYLTL

DRSSGTLSSGGEAQRIRLATQIGSRLTGVLYVLDEPSIGLHQRDNDRLINTLKEMRDLGNT
 LIVVEHDDDTMRAADYLVDVPGAGNHGGEVVSSTPNKVMKDKSLTGQYLSGKKRIEV
 PEYRREITDRKIQIKGAKSNLKNVNVDFPLSVLTVVTGVSGSGKSSLVNEILYKALAQK
 INKSKVKPGNFDEIKGIDQLDKIIDIDQSPIGRTPRNPATYTGVDIIRDVFAQTNEAK
 5 IRGYQKGRFSFNVKGGRCACKGDGIIKIEMHFLPDVYVPCVCDGKRYNRETLEVITYKG
 KNIADVLEMTVEEATHFFENIPKIKRKLQTLVDVGLGYITLGQQATTLSGGEAQRVKLAS
 ELHKRSTGRSIYILDEPTTGLHVDDISRLKVLNRIVENGDTVVIIEHNLVDVIKTADHII
 DLGPEGGEGGTIIATGTPEEIAQNKGSTGQYLKPVLERDSVE*

10 Sequence 3325

Contig_0812_pos_1152_109

is similar to (with p-value 2.0e-22)

>pir:pir|A43577|A43577 regulatory protein pfor - Clostridium
 m perfringens

15 atggatattatttttaggagtagggacttttagtactcggttcttattatcatgacgcttttc
 ttaattttcgcccatatggtaacaagggtttacaagctttatcaggggctgcttggtgcc
 acgtttttaccacaggcggttcttaagttacgcaattggtggcgattcccatgttaaat
 ttccaagaaattgggtgatctagcagggagtttaagtggatatagctggttggtatattaact
 tgtctaaagttagaagtgctccagtatattgcagtcattgttaggttttagtattacatgac
 20 tcaaaattattacctgcgtttatcgcagcggtatgttgttgcatattggaatcaagtttggt
 gagaaaaaagttccagagggactagatttaattggtgttattttattggctccagcagtt
 acatttggtcttgcaaaacttaatttctccaggggttattgcagtcacttaaacaaattggt
 agtgcaatcacttcagtaggtgataacaacccatattgcattagcagtcattttaggactt
 gttattctctgtaactgggtatgacgccattagctcaatggctacttacaagcttattaggt
 25 ttaactgggtattccaatggcaattggtgcattaaatgtacagggagcatctttgttaaat
 ggaatcttatttagcaaatataaaattggtaataaaggtaatgccttcgcgggtatttgta
 gaaccgttaactcaaatgacttaattgcaaatatccactacaactgttggtgcaaat
 gccattattggtgttgtaaatgcttgattgtcacatacagtggaactaattattgatatt
 aaaggatggcaacacctatagcaggtgctattgtactttatggctttaacgacgctgta
 30 agatctacaattacaattatcgtagtagcaattgcaagtgtagatttagcgtagcttatt
 agtgctattattaataaatttaacttgatgaatgtcggattcaagttaccacgtagaaaa
 aaccaagttaaggagagtggttaa

Sequence 3326

35 MDIILGVGTLVLVLIIMTLFLNFAPYQKQGLQALSGAACATFLPQAFLSYAIGGVFHVKF
 FQEIGDLAGSLSGIAVGILTCLELVSPVFAVIVGLVLHDSKLLPAFIAAYVAVFGIKFV
 EKKVPEGLDLIVVILLAPAVTFGLANLISPGVIAVLKQIGSAITSVGDNNPYALAVILGL
 VIPVTGMTPLSSMVLTSLLGLTGIPMAIGALTCTGASFVNGILFSKLLKIGNKGNFAVAVFV
 EPLTQIDLIKYPLQLFGANAIIGVVNACIVTYSGLIIDIKGMATPIAGAVLVYGFNDVAV
 40 RSTITIIAVAIASVILAYVISAIINKFNLNVGFKLPRRKNQVKESV*

Sequence 3327

Contig_0814_pos_2358_4124

>sp:sp|067589|SYD_AQUAE ASPARTYL-TRNA SYNTHETASE (EC 6.1.1.
 45 12) (ASPARTATE--TRNA LIGASE) (ASPRS). >gp:gp|AE000750|AE0007
 50_9 Aquifex aeolicus section 82 of 109 of the complete geno
 me. NID: g2983999.

atgaataaaagaacaacgtattgtggttttagtcacagaagaatttttaaacgaaaaagta
 acattaaaagggttggttcataacagacgagatttaggtggattaatttttggtgattta
 50 agagatcgtaggaggtattgtccaaatagtttttaactcctgacttttccgaagaagcattg
 caagttgctgaacagtagcgtcagaatattgtagtcgaagttgaaggtgtagtaacaaaa
 cgtgatgctgaaactattaaccgaaatcaaaaacagggtcaagttgaggttcaagtttca
 aatattgagattattaacaaatcagaaacacctccattttcaattaatgaagaaaatgta
 aacgttgtagaaaatattcgattaaaatatagatatttagatttacgtagacaagaatta
 55 gcgcaaaccttttaaaatgagacatcaaacacgcgttctatccgtcaatacttagataat
 aatggcttcttcgatattgaaacaccagatttaacaaaatcaaaccttgaggggtgcgcga
 gattatctagtaccttcccggtgtacacgaggggtgaattttacgcgttgccacaatcacca
 caattatttaacaactattaatgataagtggttttgataaatattatcaaattgttaaa
 tgtttccgtgatgaagacttacgtgcagatcgtaaccagaattcactcaagttgatatt

gaaatgagttttgtcgatcaagaagatatcatagccatgggtgaagatatgttacgtaag
 gttgtaaaagatgtaaaaggaatagacgtagtggtggccattcccacgtatgacatatgca
 gaggctatggaccgttttgggttcagataaacctgacactcggttcggtatggaacttatc
 aatgtgtcacagcttggttaaagaaatgaatttttaaagtttttaaagatacggtagataac
 5 aacggcgaaattaaagcaattgtcgcaaaagacgctgcaataaatatacacgtaagac
 atggatgcattaacagagttttaaataatatatgggtgcaaaaggattagcttgggttaaa
 gttgttgatgatgggtttaaagtggtggcccaattgctagatttttcgaagatgttaattgtgaa
 acacttaaacagttaacagaagctaaacctggagatttagtaattgtttagctgataaa
 cctaattgttgttgcataaagtttaggggctttaagaattaaattagcaaaagaattaggt
 10 ttaattgatgaatacaaaattaaatttcttatgggttaactgattggccgttattagagtat
 gtagaagatgcaaaacgttatgtagcagcacatcatccatttacttcacctaaaagagaa
 gatatcgaaaagctagacactgaacctgaaaatgtacaagccaacgcttatgatattgtt
 ctaaatgggttatgaacttgggtgggttctataagaatacacgatgggtgaattgcaacaa
 aaaatgtttgaagtattaggatttactaatgaacaagctcaagaacaatttgggttcttta
 15 ttagatgcttttaaatacgggtgctccacctcatggtggcatcgcgtaggtttagataga
 cttgtgatgttattaacaaatagaacaaacttgagagatacaattgcattccctaaaaca
 gcacagctacatgtcttttaactgacgctccaggagaagtatctgataaacaactccaa
 gaactctcactaagaatcagacactag

20 Sequence 3328

MNKRTTYCGLVTEEFLENEKVTLGKGVHNRDLGGLIFVDLRDRGGIVQIVFNPDSSEAL
 QVAETVRSEYVVEVEGVVTKRDAETINPKIKTGQVEVQVSNIIEINKSETPPFSINEENV
 NVDENIRLKYRYLDLRRQELAQTFKMRHQTTSRIRQYLDNNGFFDIETPVLTKSTPEGAR
 25 DYLVPSRVHEGEFYALPQSPQLFKQLLMISGFDKYQIVKCFRDEDLRADRQPEFTQVDI
 EMSFVDQEDIAMGEDMLRKVVVDKVIDVSGPFRMTYAEAMDRFGSDKPDTRFGMELI
 NVSQLGKEMNFKVFKDITVDNNGEIKAIIVAKDAANKYTRKMDALTEFVNIYGAAGLAWVK
 VVDDGLSGPIARFFEDVNVETLQTEAKPGDLVMFVADKPNVVAQSLGALRIKLAKELG
 LIDESKLNFLWTDWPLLEYDEDAKRYVAAHHPFTSPKREDIEKLDTEPENVAQANAYDIV
 30 LNGYELGGGSIRIHDGELQOKMFEVLGFTNEQAQEQFGFLDDAFKYGAPPHGGIALGLDR
 LVMLLTNRNLRDTIAFPKTASATCLLTDAPGEVSDKQLQELSLRIRH*

Sequence 3329

Contig_0814_pos_4680_5456

is similar to (with p-value 6.0e-34)

35 >sp:sp|Q57097|YGD_L_HAEIN HYPOTHETICAL PROTEIN HI0118. >pir:
 pir|C64049|C64049 molybdopterin biosynthesis protein (chlN)
 homolog - Haemophilus influenzae (strain Rd KW20) >gp:gp|U32
 698|U32698_1 Haemophilus influenzae Rd section 13 of 163 of
 the complete genome. NID: g3212178.
 40 atgaaacatcaattttcaaggaatgaattagcaataggacaagaagggtgaacttacta
 aaaaataagactgttgcagtttttaggtgttgggtggcgctcggtcatttgcagctgaggca
 ttggctcggactaatatagggcacatcatacttatagataaagatgatgtcgatattaca
 aatgtgaacaggcaaatcatgcactgacttcaactattgggtcaaagtaaagtcacgcta
 atggaagaagaatcaaattaataaaatcccgattgtaaagtaacttcttgcataatgttt
 45 tataccgagggaacatacaaaagatatcttcaataattatgatattgattattttattgat
 gcaagcgatacaatcattttataaagttcatctcatgaaagagtgttagaaagaggaatt
 gagttaatttcaagtatgggtgcagcaataagactgacccgacacgttttgaaattgca
 gatatttcaaaaacacatactgatcctatggctaaagtaattagaaatcgtttaaaacgc
 cttggtatttcgtaaaggtgttaaagtagtattttctgatgaaagtcctattgttattcgc
 50 gaggacgtaaaagaaacagtaggagataaaaatgcaatcaatagaaaagggaatgcct
 ccattctctaattgttccaagtgttaggccttatttgtgcaagctacggtgtc
 aacgatattttaaaagatacctgtgaaggcgaattaaagataaaggacaaaattaa

Sequence 3330

55 MKHQFSRNEIAIGQEGLNLLKNKTVAVLGVGGVGSFAAEALARTNIGHIILIDKDDVDIT
 NVNRQIHALTSTIGQSKVTLMEERIKLINPDCKVTSLSHMFYTEETYKDFNNYDIDYFID
 ASDTIIYKVLHMKCECLERGIELISSMGAANKTDPTRFEIADISKTHTDPMKVIRNRLKR
 LGIRKGVKVVFSDESPIVREDVKETVGDKNAINRKGMPPSSNAFVPSVVGILICASYVV
 NDILKDI PVRRIKDKGQN*

Sequence 3331

Contig_0815_pos_2100_949

is similar to (with p-value 2.0e-49)

5 >gp:gp|U59234|SPU59234_4 *Synechococcus* PCC7942 biotin carbo
xylase (accC) gene, complete cds. NID: g2661137.
atgCGTgcagaaatgaattagtagaacaacaaagaaaaacaaaagcattgtat
aaacaagagtttagcatggatgCGagcaggagcaaaggcaagaactactaaacaacaggca
cgtatcaatagatttaataactagaatcagacgttaagacgcaacatacacaagataag
10 ggtgaacttaattcttgcataattcaagggttaggtaaacaaagtatatgaattaaagaattta
tcaaaatcaattaataataaaagttttatttgaagatgtcactgaaattattcaaagtggt
agacgtataggtattgttaggacctaattggagcgggaaaaacaacattacttaattttta
agtaatgaagatcaggactatgagggtagcttaaaatcggtcagactgttaaggtagct
tattttaagcaaacagaaaagacacttgaccgtgatattagagtgattgactacctaaga
15 gaagaaagtgaatggctaaagaaaagatgggtacctcaatttcagttacacaattgtta
gaaagatttttttccgagcgtacacacggtaaaaaagtttataaaactctcaggtgga
gaacaaaaacgtctgtattttattgcgtttacttgttcataaacctaattgactcctttta
gatgaaccgactaatgatttagatactgaaacacttacgattttagaagattacattgat
gatttcggtgggtctctgtcattacggtcagtcagtcagtcgattttcttaataaaagtggtta
20 caagaatatttggtttattcatgatgggtaaaaatcgaaaaaattattggatcatttgaagat
tatgaatcttttaaaaaggaacatgaacgcccaagccatgctatctaaacaaactgaacaa
caaaataaacataagcatcaacaaaaaagaaaacaggactatcttataaagagaagtta
gaatacgaacaataatgacgcgtatagaaatgactgaaacgcgtttagaagaccttgaa
caagaaatgattaatgcaagtataatgcaagaatcaaagaacttaatgaggaaaaa
25 gagcaactgaagcaacctatgaagcagacatcacgagatggagtgaagcttgaggaaatt
aaagaacaataa

Sequence 3332

30 MRAENELVEQKQKEKQKALYKQELAWMRAGAKARTTKQQARINRFNQLESDVKTQHTQDK
GELNLAYSRLGKQVYELKNLSKSIINNKVLFEDVTEIIQSGRRIGIVGPNAGKTTLLNIL
SNEDQDYEGELKIGQTVKVAYFKQTEKTLDRDIRVIDYLREESEMAKEKDGTSSISVTQLL
ERFLFPSATHGKKVYKLSGGEQKRLYLRLLVHKPNVLLLDEPTNDLDTETLTILEDYID
DFGGSVITVSHDRYFLNKVVQEYWFHDKIEKIIIGSFEDYESFKKEHERQAMLSKQTEQ
QNKHKHQPKKKTGLSYKEKLEYETIMTRIEMTETRLEDLEQEMINASDNYARIKELNEEK
35 EQLEATYEADITRWSELEEIKEQ*

Sequence 3333

Contig_0815_pos_935_528

is similar to (with p-value 1.0e-32)

40 >pir:pir|S30712|S30712 DNA helicase - *Escherichia coli*
atgcaagagactttatcgattatttttggttataagtcatttcgaccaggacaagaagaa
attataactaaaatattaaatcatcagcatacactgggggtactacctaccgggtggcgggt
aaatcgatatgctatcaagtaccaggtttaatgcaggggtggcacaactattgtttataagt
ccacttattttctttgatgaaagaccaagttgatcaactacaagcaatgggaattcaagct
45 gcatatttgaatagtagtttgactcataaacaacaaaagagattgaagagcaaataaag
cgtggtgccattcagtttttatatgtagctccagagcgtttgaaaatacttttttcta
aatttattacgtaaaatagaaataccccttatcgcttttgagcaataa

Sequence 3334

50 MQETLSHYFGYKSFRRPQEEIITKILNHQHTLGVLPTGGGKSICYQVPGLMQGGTTIVIS
PLISLMKDQVDQLQAMGIQAAYLNSSLTHKQKQKEIEEQIKRGAIQFLYVAPERFENTFFL
NLLRKIEIPLIAFEQ*

PU3480

55

1

SEQUENCE LISTING

Sequence 3335

step.1000b10.cons.ok
TCATACCCATAATGGCAGCACCTTGCATTAATGCGTTAGCAACATTGTTAC
GTCCATCGCCAACATAAGTAAAGTTGATATCAGCATACTCTTTTTTCAATA
5 CTTCTTTAGCAGTTAAAAAGTCAGCAAGCACTTGTGTAGGGTGATCTTCAT
CGGTAAATCCATTCCATACCGGAACACCTGAATATTGCGCTAATGTTTCTA
CAGTACGTTGAGAGAAACCTCGGTACTCAATACCATCATACATACCACAA
GTACACGTGCTGTATCTTTAGCAGTTCTTTTTTACCCATTGAGAACCCTG
TAGGTCCAAGATAAGTGACGTGTGCACCTTGATCATGTGCGGCAACTTCGA
10 ATGCGCAACGTGTTCTAGTGGAATCTTTTTCAAAAAGAAGCGCGATATTTT
TATTTTTTAGCATAGGCTTTTCAGTGCCGATATATTTGGCACGCTTCAAAT
CTTCGGAGAGTGTTAATAAAAAATCTACCTCTTGTCTGTAAGTCTAATA
AAGTTAAAAAGCTTCTGTACGTAAATTTTTTCATCAATAACAACCTCTGT
GAATGAATTACAATTTTATTGTAAGCGTTTCCTGTAGTTTCAACTTAAC
15 TCAAAAATTGTAATTTTTTATTAATAAATCAATGAACTTTATTATAACATF
GAATTTGAATAGTGTGAAATTTGTCAAAAAGGTTTATGTTTAAAGTTAAAA
CCAACTTTAAAGAATTTAGAAAATTCATTAAGTGAATTTAACACATTTT
AAATAATGAAGGTAGATTAGAATCAAAATGTAAGAGGTATGTTAGTTGAAT
AAACCGTATTAACTTCTTTATATTATTAGTGAATTCCTCAAAATACTTGTAA
20 AATTATATTGATTAATTACTTGTCTTTATTATATTTAGAAACATCTCTTA
CAATAAACCACTTACATGAAATTTAAGCTAACAGTGGATTATTCTAAT
TAAATTGAAGGTCAATTCAAAGTAAAAAGGGCATATTAGTATAATAGGAGGA
ACAGTAAGATGAAAGTGTATAAGTTGAATTATCAACACCATAAAGATATTG
TTGATGATAATGTATTAACGATGTTCTGTAACAGCTGATAATCAAGATGAAG
25 TTGAAGCGTTTGCAGAAAAATTACATTACAAGATTGAACATTTATCTCCAT
TAACTAAAAAGGAATTTGAAGATGAGAAAGCGAAAGATTACACTATAGAC
TTGAACACGTGGATCACTATTTAAATTAATTATTCAATCATATATCAGAAC
ATCCTGTGATCATTCGTTTGGTTACAGGTTTTTTACGCATGAGATTAATG
TTAATGTCACTAATACAAGGAGTGATAAATGGAATAACAAATGTTAATC
30 ATATTTGTTTTTCAGTGAGTGATTTAAATACCTCTATACAATTTTATAAAG
ATATTTTACATGGTGACTTATTAGTATCAGATAGAACGACAGCATATTTAA
CTATTGGTCACTTGGATTGCACCTGAATCTAGAAAAAATATACCAAGGA
ATGAAATAAGTCATTCCTATACGCACGTTGCTTTCTCCATAGATGAAGAAG
ATTTTCAACAGTGGATTCAATGGCTTAAAGAGAATCAAGTAAATATTTTAA
35 AAGGGCGACCAAGAGACATTAAAGACAAAAAATCGATATATTTTACAGATC
TGGATGGGCATAAAATTGAATTACATACTGGAACATTAAAAGATAGAATGG
AATATTATAAATGTGAGAAGACGCATATGCAATTTTACGATGAGTTTGTAT
ATTTTAATTGATATGTAATTTTTTGAAAAAAATTTATGATTCATTTGTAAAT
AAAATTACTATATACTAAAAATTTGTAATTTTCAATTATTTAATGAGGTGA
40 ATCAATGAATAAATTATCAAAGTACATTGCAATAGCTACATTAGCATCGAC
AGTTACAATTTACGACCAATTCATACATATGCATGTGAATCTCATACTAA
AGATAACCATAATCAAACGACACAACATCAAAATGACCCCAACCTTGGTGA
ACAAAATGTAATGGCTGTCTCATGGTATCAAAATTCGCGGAAGCGAAGGC
ACTTTATCTACAGGGATACAATACTGCAAAATATCAGTTAGATGAACATAT
45 TAAAAAGGATAAAGGTAAGAAAAAATCTGCTATAGCTTTAGACTTAGATGA
AACAGTCTTTGATAATTCACCATATCAAGGATATGCTTCTATGCACGATAC
GTCTTTTCCAGAAGGATGGCATGAATGGGTGCTGCAGCGAAGGCAAAACC
TGTTTTATGGCGCAAATCATTCTTAAATATGCTGCAAAAAAAGGTATCGA
TATTTACTATATTTTCGGACCGTGATAAGGAAAAAGATTTTAAAGCTACAAA
50 GGAAAATTTAAAAAATATTGGACTACCGCAAGCGAAAGATAATCATATTTT
ACTAAAAGGGAAAAACGATAAAAGTAAAGCATCGCGACGTCAACAAGTCGA
AAAAAATCATAGTTAGTGATGTTATTTGGTGACAATTTGTTAGATTTTAC
AGATCCTAAAAATCTACTGCTAAAGAACGTGAGAACTCGTACAGAAGCA
TGCAAAAGATTTTGGCAAAAAGTATATTATTTTCCCAAATCCAATGTATGG
55 AAGTTGGGAATCGACACTTTATCATATCAATACGAAATAAGTAAGAAATGA
AAAGGATGAACCTACGAAAGTCATCAATTAAACAATTTAATCCTAAAACGGG
TGAAGTAAAAAATTCCTGACATTAATACTAAATTTCTATTAGTAAAAAGA
AAGTTTTACACATAATTATTGTATTGTGTGCAGACACTTAAATAATTAATA
AAGTTATATTTACATGGAGTACTTTAATTCGGTGTATACCATTGCGAGGA

5 TGAAACAATGTACGCGTTGTTCCATCCTTTTTTTTATAAATATTTTCAATC
GTATCTAGATTAATGTAACTTATAATATCGAAAAGTTTACAATGTGCAAT
CGAATGATTTGAAAATGCTAATTTTTTAAGAAAAAGGAGTTAAAGATTATGA
CATTAAACCTAGCTCAACGTGTGTAAATCAAGAGTCATTAACAAAAGATG
AAGCAATATCTATTTTCGAAAATGCTGAAATTGATACATTTGATTTATTAA
10 ATGAAGCCTACACAGTGAGAAAACATTACTATGGTAAAAAAGTTAAGCTTA
ATATGATATTAAATGCTAAAAGTGGTATCTGTGCAGAAGATTGTGGGTACT
GTGGGCAATCTGTAAAATGAAAGAAAAGCAACGTTATGCACTTGTGTAAC
AGGACCAAATTAAAGAAGGCGCTCAAGTGGCAACTGAAAATCAAATCGGTA
CATACTGTATTGTTATGAGTGGTAGAGGTCCCTAGTAACAGAGAAGTCGATC
15 ATATTGCGAAACAGTAGAAGATATTAAAAAGATACCCCACAACATAAGA
TTTGTGCGCTGCTTAGGATTAAACGAAAGAAGAACAGGCTAAAAAATTAAAGG
CTGCTGGTGTGATCGTTATAATCATAATTAAATACGAGTGAGCGTTATC
ACGATGAAGTAGTAACACATACATATGAGGATAGAGTGAATACGGT

15

Sequence 3336

step.1000c01.cons.ok

20 TAACATCAGTGTTTCAGTAAAAATTGGCTTCATTACTGGAGAAATGAATGTG
GTTATACTATAAAACCGGTTTCTGAGACAAATGACGTTCCGATATATGATC
AACAACAATTTATAATGTATGCGCATTTTGCTGATGAAGCTTATCAAAAA
TTGATTACATTAATTATTTTGACACATCTAGAAGAAAAATAAAAAGAGAGT
TATACGATACACGCGGATTTTAAAGTTGTGTTAGAGTCTTATCAACGGACC
AAAAAATACAAGCGGAGTACTATCTTTCTCCTCAAGGAAACGTGAAAATTG
25 AAAAGTATTATGATATTAAATTCAAATCATCCCTATGAGGCAAAAAAGATTA
TATTAAATCATTTAGGTAAAACGCACCTTTTTTAAACAATGAAACTGAGCTAG
GTGCTTTTTTTATTGAAACAATATATCAAAATAGTGACTTATTCTTTAGCG
ATCGTAACCTTGTTACATCACATATATTTAACATAGTTGCATCTTACATCC
CTGTTATTTGACAGTGTTACATAGCACTCATGTTAAAGACGTTACTGATTTAG
30 CGCATTCACCAATCAAAAATGTTTATAAAGGGGTTTTTGAACATTTACAAA
GGTATAAAGCCATCATTGTTTCAACCCAACAACAAAAGCTGACGTAATTG
AGAGAATCAGTGGAGTTATTCAGTATATGCTATTCCAGTTGGCTATTTCAT
CTATCGATATGAAAGATTACTCTAATGAAAATAAATATGTTTCACCTAAGA
AAATTATTTCTGTGCGAAGATACTCACCTGAAAAGCAATTGATACAACAAA
35 TCGAACTTATTAATAAGCTTAAAGATTCATTTCCAAATATCGAACTACATA
TGATGGTTTTTGGTAAAGAAGAACAACATTTAAAGAACGAATTCAGAAT
TAGGATTAGAAAAGCACGTAATATTAAGAGGTTTTTTAAAGATTAACTG
ATGAATATCAAGATGCATATCTAAATTTAATAACAAGTAATATGGAAGGCT
TTTCCCTAGCATTATTAGAAATGTGAGTCACATGGTGACCTTCTATCAGTT
40 ATGATATTCAATATGGTCCAGGTGAATTGATACAGGATGGTAAGAATGGAT
ATTTAGTAGAAAAGAATAATCAACACATGCTATTGAAAAAGTAAATATAT
TGCTCAACAACCCCTCAACTTCAACAGTCTTTTCTCACCATTGTATTGAAA
CTGCACAAAAGTATTCTCGAACACAAATCATGTTACTGTGGAATAATTTAC
TTCAACATTTTAACTAAAGCATAAAAATATAAGTCGTCAAATTTTGCTATG
45 AAGTCATTTATCATTAAATGGTTTCATAGTATTTTTTTGTATTAAATAAAA
GAATTTTTATTAATTCAATAAAAAAGGAACCATGAATTCATTTAAATTCAC
GGTTCCTCTATTAAATCATGTCAGATTGTGAGTTGTATATGTTTCTCTAA
AAATCATACATTTAAGTCATTTTAAATAGAAATGACCTTGTTTGCATCATC
TTTTTCACAAGATATGATAAAGCTGAAAGTTATATTATAGTTTAAATCAAC
50 ATCTTACATTTGTTGGACTAAAATCTACTTTCTATTCAAATCTCATATCTGC
TTCTTTCTACAAATTTCAATTGAACATTTTAAATTTTATTTTCTATTTT
TGCGACGTTTCCCTAATAATAACGCTCCTAAACCTGCAAAACAGAGTTCCAA
GTAACGTGCTTTAGAGCCATAATCTTCATTAGCTCCTGTATCAGGTAATT
TATCTTTTGTACTCTTATCTGAGCTATTGCCTAAATCTGAGTCGTTGTCTG
55 AGTCTGAATCACTATCGGAATCCGAATCACTGCCTGAGTCTGAATCACTAT
CAGAACTCTGAGTCACTATCCGAGTCTGAGTCACTATCCGAGTCTGAGTCGC
TATCGGAGTCTGAATCACTATCGGAATCTGAATCACTGTCTGAGTCTGAAT
CACTATCAGAATCTGAGTCACTGTCTGAGTCTGAATCACTATCGGAATCTG
AGTCGCTGTCTGAGTCTGAATCACTATCGGAATCTGAGTCTGTCTGAAT

CTGAATCACTGTCTGAATCTGAATCACTGTCTGAGTCTGAGTCGCTGTCTG
AATCCGAGTCGCTATCAGAATCCGAGTCGCTGTCTGAATCTGAGTCGCTGT
CTGAATCCGAGTCGCTATCCGAGTCGCTGAATCACTATCCGGAATCTGAGTCGC
TGTCTGAGTCACTATCGGAATCCGATTTCGTCATCATAGTATCCGTTATCTA
5 TACTAAAGTCATCATGATCAGTAATTGTTACATGAACCTTCTTCCCCATCAG
CATCCTGTTTCGTCATCATCACCAGAATCTGTTGTTGTTTGAGTCATACCTG
AAGGTTTATCAAAATGAACAATATAATTACCACTATTTAAATTATCAAATT
GATACTTTCCATTTTTCATCGGTTGTAGTTGTACTAATGATATTTCCGTTTT
CATCTTTTAAACGTCACTTTTAACTCCAGAGATTCCCTTTTTCATCATCACCTT
10 GAATACCATCTTTATTAGTGTTCATACCATACATAGTTCCCTAAGCTGTATT
TAGGTGTTTGTAGGGGAAACAATTAATGTTAAGTATTAACCAATTAACCA
GATTTATTCAGGGAATAAAAAAGCGGTAGATAATTTCTTTGGATATTCA
ATCTGGTGAATTTATCGCTTTTATGAGGACGAGTGGTAGTGGTAAACAAC
TGCACATCGCATGATTAATCGTATGATTGAGGCGACAGATGGACAGATTAT
15 GATGAATGGAAAAGATGTCCGTAATATGAATCCTGTTGAATTGCGGAGAAG
TATCCGTTATGTCATTCAACAAATTTGGTTGATGCCACATATGACTATTCG
AGAAAATATTGTTTGTAGTACCTAACTTTTAAATGGTCTAAAGAGAAGAA
AGATGAAAAGCTAAAGAACCTTATTAACCTGGTAGATTACCTGAAGAATA
TTTGGATCGTTATCCAGCTGAATTGTCAGGAGGGCAACAACAACGAATTGG
20 TGTGTGCGCGCTTTAGCAGCTGAACAAGATATTATTAATGGATGAACC
TTTCGGTGCATTAGATCCTATTACACGCGATACATTACAAGATTTAGTAAA
GGAATTACAACAAAATTAGGAAAAACATTTATTTTGTCACTCATGATAT
GGATGAGGCTATTAAATTAGCAGACAAAATATG

25

Sequence 3337

step.1000d09.cons.ok

AAACAGAAACGCGAGCTTCAATTAAAACCTCATCACCTTCTTTAGGCTCAA
AGCCTAACTTAGAAGCATTTCTCTTTAAACATCATTCCTACTGATGACACTTT
30 CTTTGTCTCTTGACATTAATAATATAAATGTCCGCTTGAATGCTTTTAAAGT
TAGATAATTCTCCTTTAATAAGAACTGACTGTAGATGAGGATCTTGGTCAA
ATTTATATTTTATATTTTAGTTAGTGCAGAAACACTTAAATATTTCAGTCA
TAGTTATCACTCATTTTAAATCATTTATATTACTTAAACACCATTAAACAAA
TTTATAATGATCATCATCACTAACTGTTTTGTGAGTTCTACAGCTTCATT
35 AACAACACTTTTCTTAGGTGTGTCGCTGTGCAATATTTCAAAAGTTGCCAT
TCTTAAATAATACGATCTGATTTTCAGTAAACGATCGATAGACCAGTCTTT
TAAATGGGGTTTAAATTGTTTCGCTCTAAACGATTGATGATCTTTGACTCC
AGTAACTAACAGTATATAAAATCAAAGTCTAAATCAGAATGATCATCTTT
AATAAATTCAATTGCTTCTTGAATTGTTAAATCTGTCTCTTTTATTTCAAG
40 TTGAAATAAAGTTTGAAGCTTGTACTCTTGCATCTTTACGACTCATATT
TTACTCCTTCAAATGTATTTATAAAATTATAGGTACAACCTTTCTAGTAAAG
TGCACCTATAATTTAAATTTATTTATTTTCTGCGACGATACCTCTGATGTG
AATATTAAATTTGCTGTGGTTCAATAGCTGTCAATTGTCGTAATTGAATTGAA
AATTGCTTCTTGAATTTGATTGTCAGTTTTAGAAATATTTACGCCATGTTT
45 TAAAGAACAAAATACATCAATGTATATTCCGCTCTCTTTAGCTTCGATTTT
TAAATCTCTGTTTAAATTTTTCGACTAATCTTCTCTAGATTTGTTTTTTT
TAGTTTCAGCAAAATGGCCTGTAATACCTTCTACTTCTGATGTCGCAATGGA
TGCGATAACAGATAATACCTTCTGGTGCTATTTCAATTTTCTCTAAATTAGA
TTGAGAATAATCTGCTACATTGACCATTCTTTATCCTCCTATTGGTTATTA
50 TCGTCCATAATATTATACCTTTCTAGGAATTTTGATTAATTTCCCACTT
CTAAAAATATGATTATTTAGAAGTCTTAAGTGAATGGAATCGTAGTGTG
ATACCTAAAACAAGATATTCACTTAAAGCACGAATGCCTGTCAATTTGAT
TCTTCACTGTAGGTTTCGTGAACATAAGTTTGGCCACCATGGAGTCATAG
TAAGGTGGTATCGTATAATTAGTATAACATGCTGATTCAATTCTCACTCCA
55 AAACCGCCTGGAGCAAGATATTGGGTAATCTTGCTGGTGATGGCATAAAG
TTTTTTGAAGGATTTTCAGCATGATTTCGAAATTCATAGCGTGACCGTTA
ATGGAAATATCTTCTTGTTTAAAAGGTAACGCCCTCACCCATAGCAACTTTG
AGTTGTAATTTTACTAAATCTACTCCTGTTACCATTTTCAGTTACTGGGTGT
TCAACTTGAATACGTGTATTCAATTTCCATGAAATAAAATTGGTTATCATCT

AAATCATATATAAAATTC AATTGTACCTGCGTTTTTCATAATTTACAGCTTTT
 GCGGCTCTAAATGTCAGCATTACCCATTTCTTGGCGTTTATCTTCACTTAAA
 ACTGGTGAGGGTGCTTCTTCAACGAGCTTTTGCATTCTTCTTTGAATTGTA
 CAATCACGTTACCTTAAATGTATAACGTTTCCATAAGTATCGCCAATAATT
 5 TGAATCTCTATATGTCTAAAGTTTCTATAAAATTTTCTAAGTATAAACCA
 CCATTTCCGAACGCGGTTTTCAGCTTCTTGTGTGTCATACGGTAACCAGTT
 TCAAGTTCTTCTCATCACGAGCAACCCGAATACCTTTTCCACCACCACCT
 GCTGTGGCTTTGATGATAACTGGATAGCCGATTTTTTTAGCTATTTTTTTA
 10 GCGTCATCTATACTTTGAATAAGTCCTTCACTTCTTGGTACTACAGGTACA
 TTGGCTTTAATCATTTTCAGCTTTAGCAATATCTTTTATTCCCATTTTCTGA
 ATTGATTGATAGCTTTGGCCCAATAAAATTTAATTGGCATGCTTCACACAAT
 TCAGCAAAAATCACCATTTTTCAGCCAAAAAGCCGTATCCAGGATGAACACCA
 TCACATCCTGTAGAAGTAGCGATAGATAAAATATTTGGTATATTTAAATAA
 GAATCTTTTGATAAAGTAGGACCGACGCAGTATGCTTCATCAGCAATTTGA
 15 GTATGTAATGCATCTTTGTACCTTCTCCGTGTACAAATTATGTTACAAG
 TTTGCATGAAGGTTTATTTAACATCGCTGCAGCTGTAGGTGTTTATAGTCC
 AACGGAGATTACTTCCGACCATATTTATCTATAGACAATTAGATGGCACTAC
 AACGTCCTATTGAGATTATAAACTTAAATTAATTTCTTAAATCACAACTT
 GAAAAGGAGCGAGACAGAATGCTGTCTCACTCCTTTTCAATTTAATTATCT
 20 TCATTCATTTTACTTTTATCTTTTAAACCATTACGTTTGAGTTGTTTAAACA
 ACCCTATCCATACTCACATGGTCACCTTTTTTGCATCTGATGATATGAAA
 CCTAAATTTTCTTAAAGCTGATCTATATCCGCATCTTTATAATCAATATCG
 ATATGTTTCGATGGCCTTTTCGATCTTTAAATCAACTTTATATGTGACCCCG
 TCTATCCCTTTGAATGCTTGTTCATCATTTTTTAAACATTTGTTTAGCTTCA
 25 TCTTTATCGATGCTTAAATCATCATATTTAATTGTTGCTATAGTAGATTGT
 TTGAGAACTTTCATCATCTTTGTATGTCAAAGATGTAATCACTTGTTTACCG
 TTGACTTCGCCCTTCATAAGTTTTTGTGCGTTCTTTCCACATGCCGAAAGA
 GATACAGCACTTATACTAATACTACTAAAAGCGATATGACTTTTTTCATA
 GTAGTCGCCCTTCTCTCATCTTTAATCATGTTAATTATATCATCTTTAATTT
 30 GAATATGTTACTTATATGAAGATAAATATCTTTTATTTTGAAAAATTTAAA
 CATCTATAAACATGGAATTCAGAATTTCACTACTTAGAATGTTATACTACT
 TTTAAAGACTTATTAATGGAGGAATGACTATGCCAGAATCTAAAGCGTTAC
 TTAAATCTTTAACAGATGTTAATGGTATTTCTGGACATGAAATGCAAGTTA
 AATCT

35

Sequence 3338

step.1000f04.cons.ok

TAATATCTTTTTTATTTGGTAACCAAACGTTTTTCAGCTTGAGTAAATGGAT
 40 AAATTGTATCTGATGCGGCTACTCGAGCTATTGGAGCTTCTAATGAAAGAA
 TTGCTCGCTCTGCTAATTCGCTGCCACTTGTGCACCCACACCAGCTTGAC
 GTTGTGCTTCTTGTACAACCTACAGCACGTCCAGTTTCTCAACTGATGCTA
 CTAAAGTATCTATATCAATTTGGTTGTACAGTACGTAAGTCAATAAAGTTCAA
 CTGAATAACCATCTTTTCTAACTCTTCAGCAGCTTTTAGTGATTCTTTGTA
 45 CCATTGCCCCGTAAGATATTAGAGTAATATCATTACCTTCTTTTAAACAT
 TGGCTTTTCCAATGTCAATTTGTATTCTTCTTCAGGAACCTCTTCACGGA
 AAGAACGATATAATTTTCATATGTTCTAGATATACAACCTGGATCATTACTTT
 GAATAGAAGAAATTAATAATCCTTTAGCATCATAAGGACCTGATGGAATAA
 CTACTTTCAAACCAGGTGATTGAGCTAAGATACCTTCTAAATTTATCAGCAT
 50 GCAACTCTGGAGTGTGGACGCCACCACCAAAAGGTGTACGAATTTGTAACAG
 GCGCTGGTTTAGTTCACCTGAACGGAACGAGTACGAGCAATTTGACCAG
 CTACTTCGTCAAATCTTCAATAACGAATCCTAAGAATTGAATTTCCATAA
 CAGGACGGAAGCCAGTCACTGCTAAGCCTAGTGCAAGCCACCAATTTCCAG
 ACTCTGCTAATGGTGTATCAATACTCGATCTTCGCCAAATTTCTTTTGTA
 55 AACCTTCAGTAACACGGAATACACCACCGTTAACACCAACGCTTTCACCGA
 AAATAAAACGCTTTCGCTCTTTTGAGTTCACTTTTAAGCGCATCGTTAA
 TCGCTTGAACCATTTGTCATTTGTGCCATGGCTTACTTCGACTCCTTCTCTT
 TGTAATTTTCATATTGTTCTGCTAAATTTTGAGGCATTTCTTCATACATGA
 TATCCATTAGAGAAGTAACAGTTTGTTTTTCTGTATTGTCAGCCTCTTTAA

TAGCTGCTTTTATTTTCAGATTTTGCACGTTCAACCACTTCATTTTCTTTGT
 CTTTCATTCCAAAGACCTTTAGCTTCTAAATATTTTCTGAAACGTACTAATG
 GGTCTTTTCTCCCATTCAGCATCTTCATCTGAAGTTCTATAACGAGTAG
 GATCATCACCAGCCATAGTATGTGGTCCATAACGATAAGTTAAAGTTTCGA
 5 TAACAGTAGGACCTTCTCCTGCTACTGCACGTTACGTCGTTCTAATGTTG
 CTTGATAAACAGCTAAAGCATCCATACCATCAACTTGAATTCAGGGATAC
 CAACTGAAATAGCCTTTTGTGCTAATGTTTCTGCAGCTGTTTGTCTACTAC
 GTGGTGTAGAGATGGCATAGTTATGTTTGAATTACAAAAATTGCAGGTG
 CTTTGTATGCAGATGCAAAGTTAATTCCTTCATAGAAGTCACCTTGTGATG
 10 AACCACCATCACCTGTATAAGTAATTGCGACTGCATTTTGGCACGTTT
 TAAGTCCAAACGCTACACCGGCAGTTTGAATATATTGTGCACCGATAATAA
 TTTGAGGGCTAAGCTTACATTAACCTCCCTCAGGGAATTGGTTACCTTTGAAGT
 GTCTCTTGAGAATAAGAATGCGTCTGTAAGAGGTAAGCCATGCCAAATAA
 TCTGAGGCACATCAGGATAACCAGGTAAGTGAAGTCTTCACTTTCTAAAG
 15 CATACTGAGATGCTAATTGTGAAGCTTCTTGTCTGCTGTTGGTGCATAGA
 AACCTAAACGTCCTTGTCTATTTAACGAAATAGAACGTTGATCAAGAATTC
 TAGTCCATACCATTCTTTCCATTAAATCCACTAATTTGTCGTCTGTTAAAT
 CAGGTACTAAGCTTTCATTAAACGACATTTCCGTCCTCATCCAAATTTGAA
 CCATTTCAAATTTTCGATTGAGTCTCATTTAAACCTTAACTGCATCGAATT
 20 GGGCTTGTAACCTTAGGAGCCATTCAATTCACCATACCTTTCCCATGTAAAT
 AGAAATTCATTTTATCCTCCTTAATTGTATCACAAATAAAATCTCTGTTA
 AACGATTTCAAAAAGTTCTGTTCCACATAAAATAATACAGTTACATGACC
 TGTGTATTAATAATCACGTAACGTATTAACCCATTATTTATTTAATTG
 GTCTACATCTTGTTTTCTCTTTTACTTTCTCAATGGCTTTTGAGTAAGC
 25 ATTAATTTTATTTATTCATTTCTTTATATGCTTTAGAAAGATCTTTCGATTT
 TCCGCTACTTCCGATTGTGTTGCATTATCTTCATTCAAATAAGAAAACAG
 TTCTTTTCTCTTATTAAGTGCTTTTGTGAAGCCTTGCATAAGCATCATG
 TGATTTATATTTTCTTTAATAGCACTATCAAGTTGTTCAACTTCTTTT
 CTTTGCTTTGTTTCTACATGTTCAAGATATTGCTTGGCTTGTTTAAATGC
 30 TTTTTCAGAAATATCTAGAGCCTTCTCTTCTTTTCAAATTCCTTTTGTCT
 TTGTTTTACATTTTCAACTATATCTTCAGCTGCTTTTACGTGTGCTTTG
 ATCTTTATCATTTTACCTTTTTAAATAATTTTGTCTTTTCTCTTAATTC
 GTTTATTTTCTTACTTACTTGATTAAAGGTTTCTCTTGGTCCATAGCTTT
 TTGTACTTGATCATCATATTTTAAATTTCTTTTATCTATAGTACATCC
 35 AGCTAATAACACTGCTGACGTTAGGACAATTGCTACAGTCTTTTTAAATTC
 CATGAACGTGTTACCTCCTCAAAGTTGTTGCTTGTAAATATTAATAATAAA
 TAATAATAATCAATATAAAGCTATCATTTCTATATTCTAGTGCAACATTT
 AAAGCAATAAAATCTTTGAGTATCTATTAATGTACTTGTGCTATCTTATC
 TGATATATTTAATAAGGAAGGTGTATGACATGATAACAATGAAAGATA
 40 TTATAAGAGATGGTCATCCAACACTTCGTGAAAAAGCGAAAGAATTAAGCT
 TCCCACTTTTCAACAATGATAAAGAAACATTGCGCGCAATGCGTGAAATTC
 TAATCAATAGTCAGGATGAAGAAACCGCAAAACGTTATGGTTTACGTTCTG
 GCGTAGGTTTAGCTGCTCCACAAATTAATGAACCAAAACGTATGATTGCTG
 TCTACTTACCTGATGATGGAACGGTAAATCGTATGATTATATGCTCGTAA
 45 ATCCTAAAAATAATGAGTTACAGTGTACAAGAAGCTTATTTACCAACTGGCG
 AAGGTTGTCTAAGTGTTGATGAAAACATCCAGGTTTGTGTCATCGTCATC
 ATAGAGTCACTATTAAGCTCAAGATATTGATGGAATGATGTTAAATTAC
 GTCTCAAAGGTATCCTGCAATTGTATTTCAACACGAAATTGATCATCTAA
 ATGGCATTATGTTTTATGATTATATTGATGCCAATGAACCTCTAAACCAC
 50 ATGAAGAGGCCGTAGAAGTCTAAATCATGAATTATTAATAAACACTTCATT
 AAGTTGTATTTTATACTTACTTAATGAAGTGTTTTTTTAAAGTAATTCAA
 ATGTCATTAGGATTAATTT

55 Sequence 3339

step.1000f07.cons.ok

TCAATATTATTTTCTTCAATATCATTTTATTAATATAATTAGGTAGAAG
 TCTACGCTGCTCTAAGTGGTAGATGAATGATCTCTATTTTTATTATGTT
 ATACATTTATATTATTTCTGTAAGGTTCTGATTCTAAAAGATATTTATAA

TAGGAGTCTTTCTTTTGTAAATCTGTTTCCCCCCACACCTTATATTTCTTTA
TCATCTGTCCATGGTGTGGTGTACGTTATACTCATACTCACTATATGGT
TGATCTTTAAAAACAATATCTTTTTCAGCATATCCCATTATAATATTA
5 CTCATTTTTCATTTTATTTTATTTTATCCCAACCCCTACGTTCTATAACG
TGATTGATGATTTTCTCACTCTTGTGTTTCTGAATATATTGTAGTATAAAG
AAAATAAAGTTTATAAAATAGAGCTTGATGATATGAATGACAATAAATAGA
CAATAAATAACCGCATCATAATCTAAAATGATTATAATACGGTAAATTTTG
ATATTATTTTATAACTCAATTGTAGCTGGCGTATCTAGTTCATCATCATTC
ATTAATTTAATTGAACGAGGTATAAATTTGAGTACACATAATTCAGGATCT
10 TCTTTTGAATGAAAAATGTTTGTCTTGAGTTTCCCATACCAATCAATA
ACTTTTGTGATTTTGCACAACCTTCTATAGTGGCATCTATTTCAACAAAGCTA
TTGTTTGTGTTTCTCGTTATAGCCTAATAAAATATGAGCATCAGGATTATTT
TCTATTTCTTCGACTTTTAGTGAATTGATATTCGTTTGTATATAAGTTT
AAGTCATCGTTGTAAAAGACCATATATCTGCTATTAGGTTTATTATGATGT
15 GCTGTTGATAGGACACCAATTTTGAAGAATTAATACTTTTTCTATTGCT
TTTGTCACTTGTGTTTATTCAAAATATCATCTCCTTATTATTGTATATA
TCCTTATCACACATAATTTAAACGAAACAACCTTATGATGATTTTAAAGAA
GGTTTAAACAAATATATTTGACTTTCATCGTGTTTTAAGGTGACTTGGAAA
GGAAGATGATGGTTCATGTTGTATAGCTAATTCGTTTGTGTTGATTTGAT
20 TTATCAAGTTCATCAAGAATACTACGTTCAATATAAAATGTTGTTTAGTT
TGTGGTAGAAGGTTTGAATTGCGCCATGCTCATGATTAAACGTTTAAATA
ATAATTAATGAGTTGGTTGATAATGTATTTTTAAAAACGTAGCGCTGTTGA
CTATCAGATAGATAACGATCATTATTTTATTAATAATAAAAGTGATAG
TTACCATCTTTGCGACTCATTAATAAATTTGTTAGCTATCTCTAAAGGTTCA
25 TTTAAAAATGGTTTCGATTAAATTTATAGAGGTGCATGAGAGGTAGGGGACTA
CCATAACGATTAAATAAGGCAATAAAAGGACTCTCATCTTCTATCAGTTGA
TAACCTACACCACCCCGTAACGCATTAATGCGACTAAATGACAAACAAGG
TCAGATAAATTGAAGTGGATAAGTACCATTGTTGTATATTTAAACACTTT
AAAGATAAATTTATTGTAAACAAATTTAGTTGAAGGTATACCGGAATCTAAT
30 ATAAATGTTTATAATTTTCAAAGTGTGTGGACGAATGTAAGTACGATTTA
CGTTTGGTAAATTAATGAGTTTGTGTTTCTAGATTTTCAATGCTATAAAA
TAGTAATCGAAATTTAATCGCTTCATCATATCATAAGTTCGATCAATAGAA
GAATTTTGGTGTAGTAATCCTTCAACAGTTATACTATAACGGATAGCTTTA
TTTTTATTTTAAATTTTAAATGGCATAAATGTATATCGTTCACTGACATA
35 TTCTTAGTTTCAAAGACTAACATAAATTTAACCATTTTTTGTTCATTTTA
TAATCTTGGTGGCTATTAAGAAATAATAAATTTATTTCTTTGATAGATTCA
AATTCATTTGTAGATGTTAATCTCATGGCTAAACCTATATTTTTTTCGAAT
AATTTTTCAAAGCATCGATTAAATAAATTAAGTTACCTCTCGCGTCGAA
ATATCAGTGATATCATTGATAAAAAATTACAGCTTGTGGCAGGGAGGTAAAA
40 TCAATGTTGTAATATTCATTAAAAACAAATTGAAATAATTCGTTGAAATTT
TGAAAACGTATAAACGCTTTAGAGTTTTTAGTCTGATCTTGAGGATAAAAT
TCATTGATATCTCTTTCAGTCGTTTCAGTAGCCAAATGATCACTAAACTCA
AATTGATTAATTAATCAATATATTGTGAAAAGTCAGTATCACTAAACTT
ATCGAAGGTAAGGAGTCTAATCTAGAGATAATCGTTTCGTTATTGTTTGGG
45 CTACAACCTAAATAATTTTTAAATTTGATTTGTAAAATTTGGTATGACTACTA
AATCCAGATTGTTCTGAAATAGCGTTAATTGAGTCCTCAGATGAAAGTAGC
CTTTTTATAGAGTCAATCACTTTTAAAGCTTGTAAAATAATCTTTAAATTC
ATATTAAGATATCTTGCAATAGGTTAGAACAATAAGATTTCAGATATATTA
CAATGTTGAGCTACTTCTCGTAGTGAAGTGACTGTGATACTTTGAATGA
50 ATAAATGTTAATCCTTCAGTAAACACTGAATGATTAACTGCAATTTGAGGA
ATATATTTTTCTTATATCGAATGACTGCTTCTTTTAGTAGCGTTGTATT
ATTTTAGAGATGGATTGCTCATCTTGATTTTCTCCTTTTATTAATGTTGA
ATAGCTTGTAATAAATGTTTAAACAAACTGCTTGATTGTAATAAATGT
CTGTCAAAATAACATTTAAAAATTTATTATCTTTACTATAAAAAATAATA
55 ACTGGAATAGAGAGTAATACGAGATTGTAAGCACTATTAATAAAATAGATG
TCAGATTGATTAATGATAGCTATATGATTACCTAAATCTTTTGTTCACCA
TTAATAGTGATTTTAAAGTCATTTGTAAGTGAACATCAGAATAATCTTA
TTAATATTGCGTTTCAGGGGAGAATGATATCTTTATATATATTTAAAAAA
CATGTCATTGTATATTACATTAGCTCCTTGAATGAATTGAATAACTAAAT

TAAATGTGACCGTGT

Sequence 3340

5 step.1000f11.cons.ok

CCATAACCATGATCAATAACTAGTGGATAATATTTAATCCGTCCAGAATGT
GCAAAATAACGTTTCATCTTCAACTATCAAATGATTAACCTTACTCATTGAG
TTCCCTCCCAATAGTGACTCTTTTAAATATAGCTATCTACATATGAAGCTA
TAAAAACTATCAACTGCTATCGTTTAAAAATAAATCAATATTAATAGGATA
10 GTAATCAATATACTAATAATTTTCAGATTTGTAAATATTTCTGATAATTT
AGAATACTTACGTTCCGAATCTTTTGTGATAAAGAAATATATGTCATTGAG
GCTTAGAATATATGTTTAAAGGGAATTTAGAATAAAATATTTTACGGAGGA
GATTAATATGCATTGGTTAAAAATATTTTATCATTTATATGCGCAACCAC
GATTAGCGTGATATTACTTATTATAACTATATTAATGGATGCGTTACTACA
15 AAACACACACTTAACTCAGTTATTACTCAATATTGATTTTTTAATCAACCC
TGATGAAGTGCCAACAATTATTGAAGTACTGATTCAATTAAGTATTGGAAT
ATTGATTTATCTCGCCTTTTTAATTATCTATCATTATTCAAAATCCTTGTA
TCATCTAGCATACTTACCCTTTAGTATGTGATATTTACTTTGATGTATCCACT
TCTCGTTTTTCTTGCGCAACGTCCATTTTTTTCCTTTAGTTGGAACGAATT
20 TGCATGGTGGTGTAGTTGCACATCTTTTTTTCATCATTTAATGGCGACTTG
TCTACCTATCATTTTCGAAAAAAATTTTATGATTTTAAACTCATATCTCAA
TGTCTTTTACACATTATAAAATAAATATTGGCAGTAGAAAACATCATAGATA
ATTAGACTATGATATCTCTACTGCCTTATTAGTTATCTGGCTCTCCATCT
TCATCATACTTCGTTACTTCTCCATCTTTATCGATAATATAAGATCCTTCT
25 AAATGGCCTTCTTTTATCTGTAAATGAGAACCCCCAACTACCATCGCTATCT
TTTTCTGGTTCTTTGTAAATATATGTGTCTGTATCTAATTGATGCCCTTCG
TAGTCTTCAACTATATCGATAACATTACTTCGTGTCACCGTTGCTTCATCA
GAACTACTTGATTGATCTTCAGATTCAGAAGTATCATCTTCATCTGTAAAA
TTATCACCAACTGAAAATTCAGTCCTCATCATCTGACTTATCTTATTGATT
30 TCACTTTCACTAGCGTTGTATAACTTAACTGTACGAGCATCATCTATAATG
CGTTGAGATTCTCTTTTAGAGTAATCAGCGTCACGCCAACTCCCTTGAAAA
TGTGATGGTACACTATATATTGTAATTGTGCCATCTCCATTACTTTTATAA
TAAACTGAACCTGCTGCAGTAACACTTGCTGTTAACAATTCTGTTCCCTCA
GGTAGTTTGGCTGATTTCGCTCGGGTGATAGGGATCTAAACTTTATTAGAA
35 ATATTTTGTAGTACAATTTCTAAATCGTCAAAAGGAAGGTCACTTTCACCT
CTATAACTATCTAAAGCAGTTAACCAAACCTTGAGCATAGTATCTATTTCGAA
TGGGTATCTTGTGTGTCTGCTGATTATCTTCTGTTTGATTCTACTT
ATACTTGATTTTGAATCTCTGTTACTTTTATATCCGATTGTTGATCATCA
TTTATCGTTGAGCTTCCATTATTATCACTATCACCAGACCAACCAGCT
40 AATATTAATATTGTAAACGCACATATCACTGCTAATTTCTTCAATTCAAAT
ACTCCCTCACATTCTAATATAATTAACTCAACAAAATAAACCACTTTATAA
AATAATTATAAACTAACTATTTTAAAAATGTACAACAATTTAACAAATTC
TTAAAGATTGCTATTTACGAAGCAATACTTGCCGCTCTGTTTTACTGCTAT
ATCATTAGCAATGGTTTATGATTATTAAAGATTTTAAAGTATCTTAAATTT
45 TGTATTTGAATTCTTTTAACTACTATCATCTCGCAATCAAGGTGCTAAAC
TATAATCGTAAAGTAATTTAAATACATCATATATCATCAATTCGAAATGGG
GTTATCAATGTGAAAAAGATTTCTCTAATTGCAACGACTGTATTAACAGGA
TTATTATTATTTCCAAGTGTCAATGACACAACGACACATGCAGCCGAAGTA
ACATCTCATGATGCAAAAGTGCTTTAAATGAAGCTCAGCAGTCATTGTCAC
50 AATATGACAGACAACCTCAATGACAGCATGAATGCATCATTTGATGGTAAAA
TTAACATTAATAATGATTTCAGATGTAGGCGAAGGGCAACCTTTTGCAT
TAATTTCTTCAAATCCTCAAATTAACGCAACTATCACAGAGTTTGATATTA
ATAAAATTAAAGAAGGCGATGAAGTAAATGTCACGTGAAATAGCACAGGTA
AAAAAGGAAAAGGAAAAATTTTAAATAGATGAACTTCCTACAAGCTATG
55 ATACAAGTGACGATAGTACAGCATCATCGGCACAAGCAGGGGCACAAGGTG
ATAGTGAAGAAGGAACGAAATGACGACATCTAATCCTACAATTAATCAGC
CAACAGGTGGTAAAAGTGGCGAAACATCAAAATATAAAGTTATCATTTGGTG
ATTTAGATATACCCGTGAGATCAGGCTTCTCTATGGATGCTAAAAATCCCTC
TAAAAACTAAAAAGCTACCAAATAACGTGTTAACAAAAGATAATAACGTAT

TTGTCGTTGATAAAAAATAATAAAGTTTCAAAAACGTGAAATTAAAAATTGAAC
 GTAATAATGGTGAAATCATTTGTGAAAAAAGGATTGAAATCTGGCGATAAAG
 TCCTTAAAGTCCAAAAGGTAAATTTAAATGATGGAGAAAAAGTAGAGGTGT
 CATCATGATTGAATTGAGAAAAATCAATCGACACTTTAAAAATGGAAATGA
 5 ATCAAATCATATTTTGAAGACATAGATATTCATATTGATGAAGGTGAATT
 TATTGCTATCATGGGTCCGCTCGGTTCAAGTAAAAGTACGTTAATTAATAT
 CTTGGGATTTATTGATCGTGGATATGAAGGAGAACTTTTTTAACAATGA
 GAATTATCAAAAAAGCTCAGATAATAAGCTCGCAGAAATTCGAAATCATAC
 TGTAGGTTTCGCTCTTTCAGAAATTTAAATTGATTCAAAAATAACACTATTTT
 10 AGAAAAATGTAAGTATTCGCTAAATTTATAATGGTTTGAGTAATAAAGCGAG
 AAAAAGCAAAGTTTTAGACGGACTTCATGACGTCGGTCTCAAAGGTAAAGA
 GAATCTCTTACCAAATAAATTTGTCAGGAGGGCAGCAACAACGTGTAGCTAT
 TGCCCGGGCAATTATTAATGACCCTAAGTTTATTATTGCCGATGAACCTAC
 TGGCGCTCTTGATTGCAAGACATCTCAAGATATTATGGAGCTTTTCGTAAA
 15 ACTGAACAAAGAACAGAATACAACTATTATTATGTAACGCATGATCGTGA
 AGTGGCGGAAAAAGCTGATCGTATTATCCATATTTTGGATGGGCGTGTGCA
 AGAGGAAGAGGTGTTGACACGTGAATAACTTTTCAAATGTCATATCAGTTG
 CTATTCGTTCCATTTTAAAAACAAACG

20

Sequence 3341

step.1000f12.cons.ok

AAAATTAGGTCCAGAAAGTTTCCCATCTTTCGAAATATCGGTATAAATCAC
 ACCACCCAAAGGTAAATGCTCAATTTTGGCAACATAATCAAATAAATTGAG
 25 TTTAGCATCCTCTTTCCATCCATTAATCTTTATTTTCTCTCCAAAAGCATC
 TACGGATAAGTAGAGTTTATTTGGAAATTGATGTGTCATATGTGTTAACCA
 CTCTATATCTTGGATACCTTTTGTACCTACAATACAATAGTCTATTCCTGA
 ATGAATATAATTTTCAATTGTTTGTGTTTGAACGAATGCCGCCACCCACTTC
 TATAGGCTTAGTGGTCACCTTTCTTAAGGAACGGATGTAGTGAATTCCTTT
 30 TACCTCTTTTGTCTTTAGCCCCAATTAAATCTACAATATGTATACGTTTCAC
 ACATTTAAATTGACTATAAAATCGGATACTGTCTTCTACAGATTTTTCAT
 TTTTCTTTTGAATCATATTTGCCTTCTGTTAATCGAACACTTGTGTAATT
 AATTAAATCAATAGCTGGCCATAAATCAATCATTAATAAACCCCTCCTTTAA
 GCGCTTGATTTAGAATCTCTAATCCATACGTTCCACTTTTTTCAGGATGAA
 35 ACTGGATACCTATATAAATTTCCGGTATTGAATGACTCCCGGAATCTTTGTAC
 CATAGTCAGCATAAGCTACTACATATTCTGACATTTCTGCTTGATATGAAT
 GAACAAAATACACATCACTTTGCAGTAAGGGATGTGTACTCTTTAATTCAT
 TCCAACCCAAATGAGGAATAGGATGAGATGATTGGATTGGCACTATATTTTC
 CCGGGACAAGTTCCAATCCACTAACGTCACCTTCTGCGCTATGTTGAAAAA
 40 GTAATTGCATACCTAAACATATTCCAATTATCGGTTTATCATGTATATTTT
 TAAGCATATCTTTAATGCTTTTTCCTTCTATAGAATGCATCGCATCCTGAA
 AATGTCACACCTGGAAGTACGATAGCTTCAGCTTTTTGACGTCCTTTAT
 CATCACATGTCAAAATCACATCATATCCTAAATGTTGAATTGCTCGAGTGA
 CGTTACTTATGTTTCCCAATCCATAATCAATAATCGCAATCATTCAATTAC
 45 TCCTTTAGACGATGGAATACGTCCATCTTCATTTTGTGCAAGAGAAATCTT
 TAATGCTCTTGCAAAAGATTAAATATTGCCTCAATCTCATGATGTGTATT
 TCCACCTCTTAATAAGTCAATGTGAACGGTTAATCGCGCATTAAATTATCAA
 TGCTCTAAAAAATTCTTCAACTAGTTTCAGTGTCAAAAAGTTCCCTACCTTTTG
 AGCGCTCAACTTGCTATTAAATGAGAAATATGGACGACCACCTAATGTCCAC
 50 TACTGTTTCGAGCAAGCGCCTCATCCATGGGTACATATGAGCAACCATATCT
 TGTAAACTTTGTTGAGTCTTTATTAATTCAAGAAGTAATTGTCCAATAAC
 TATACCTATATCTTTCAGTTATATGATGATCATCAACATACGTATCTCCAGT
 GGCTCAATAGATAAAGTTAATCCACTATGAAAAGTAAATAGCGTTAACAT
 ATGATCTAAAAATCCTACACCAGTATTAATATGAGATTGTGTTCCGTTATT
 55 AGCCAGTGAAATATTAAGCTGAGTTTCTTTCAGTGTTCCTTTTGATTGATA
 GTTCATAACATACGCTCCATTTCTTCACATTTTTCAGTAGAAGAAGTTG
 TTCTTCTGTAGCAATAGAGTAACGAACATACGCTTTTCATTTTATTAAACAG
 ATATCGTAATAACCTATCTTTATCGGCGTACGTGAATCGATTTTTCATGAT
 TAAATAAAGATTCTATATGTGCAATATGCTCAGCTGATTCTGCCACTTCGT

TAAATGTTTTTGTGATAAATGAATGACTGAATGTCCTGTCATAAAGTCAT
 TCACAGAAAGCCCATTTGTAAATCTAGCTGTTTGATTTAGTAGGAAGTACAT
 GACTTGGACCTGCCACATAGTCCCAATAAATTCAGGAGAAAAATGACCTA
 AAAAAAGTGCCCTACATATTTAACTTTATCTATATACATTTTCAGGTGCTC
 5 GAGTTTGAATAGATGCATGTTTCAGGTGCAATCGTGTTCATAATTAAACATG
 CTTCTTCAGTATCTTGAGCTAATATTAGATAATGATGGTTAGCAATACTTT
 GTGATATGATATCTTGCCGTTCAACATACTGAAGTTCTCTTGATTATAG
 TGTTCAATTTGATTAAAGTACCTTTTCATTTTCACTAATCACATAAGTACAAG
 CCATTTTCATCATGTTCTGCTTGTGCAAAATACGTCATAAGCGATTGCGTCTA
 10 AGTCAGCACTTTCGTCTATAATCAAGGCTATTTCTGTCGGTTCTGCGATT
 GGTCATAGCCTGTTTACCGAATACAAACTTTTTGBCATAAGCAACATATT
 GATTCCCTGGACCTACGATTTTGTGCGACTTTTTTATAGTTTCCGTGCCAT
 AAGTTAGCGCCGCAATACTTTGTGCTCCACCGACTTGATAAACATGATGAA
 CGCCTGTAATGTAACAAGCGGCTAACACCTCTTGACATATAACCGCTATTTT
 15 GAGGTGGGGTAAACAACAGTAATCTCATTAACACCTGCTACTTGAGCAAGTG
 TTGCAGTCATTAATACTGTAGACGGATAGCTAGCCTTACCTCCCGGCACAT
 AAATACCTACACGTTTCGATAGGATGGTATCGTTTCATAACATTCAGTTTGT
 GAGATGATTCTGTTTACCTTAATATTTTCTTGGTACACTTTAATTCTCT
 GATAGCTTTGCTCTAATGCATCTCGTGTTCATTATCTAGCATGTGCTATG
 20 CATTTTTTAGTTGGCTTTGCTCTAATTCAAGCTTCTCAGTTTCCACTTGAT
 CAAATTGTAAATATAAATTTTTTAAAGCTTTATCTCCCTGTAATTTAACTT
 CTTCAAAATCTGACTCACTATCTCGTACAAAGATTCACTAGAGATTCAA
 CATTATTAATTTCTTTTAAAAATTTGTTGTGCGCTAAGCATAGTTAATAGAC
 ACTCCTAAGTGTGATTAGTCTCTCTATTTTCAGATGATTGCTTAAAAATAT
 25 GATTCTTTATTTGTAATTAACTTAGCGTTAATTTCACTGATATGC

Sequence 3342

step.1000g04.cons.ok

30 ACACCATCCCTATCATATATACGGCAGTAAAGATGAAGACCCAAACAACAT
 CTAAGAAGTGCCAGTATAAACTTACTATAAAATAATTTAGGAGCATTGTATG
 AATCCAAACCACGAGTGCCGATTTGAATTAACAAACAAATAACCCAAACAA
 TACCTAATGATACGTGTGCACCGTGCGTACCTAGTAGTATAAAGAACTAG
 ACCAGAAGGAGCCAATAGTTGGGTAAACACCTTCAGAAGCATAGTGTGCGA
 35 ATTCGTAAATTTTCGAAACCTACGAATACAAGACCTAGGATAACTGTGATAA
 TCATCCAAACATCATTAAGTTTGTGTTTCTTGTGCGCATGTAATAAATTG
 CAATACCACAAGTATAAGAACTAATTAATAATGCAAAAGTCATTATTAAAA
 TCAAATGTAAATTCGAATAAGTCGGTAGTTAATTTACCGCCATATCCGCCAC
 CATGTTGTAACGTTAATAACGTTGCAAAATAGGGTACCGAATAACGCAAAAT
 40 CAGCTGTAAGGAAAATCCAAAAGCCAAGTTTATTTAATTTCGCCTTCGTGTG
 TACGAGAATCAATAGTATTTGCATCATGACTCATGACTTACAGCCTCCCTT
 TCTTTAATTCGAGCTTCTCTTAAACGAGCTTCAGTTTCTGCAACTTCTGAA
 GCAGGGATGTGGTAACCATGATCAATTTGGAAACTTCTCCAAATCATAGTA
 ATGAAGATACCTGCTAAACAGATAAGTGCTGGAACAATAGATTGGAAGATT
 45 AAGAAGAAACCACCAATAGTCATAAATATACCCATCCAGAATCCTACTGGA
 GTATTGTTTGGCATATGAATATCTTTGTAGTTATGGTTGTCTAAATAATGA
 CGACCATGTTCTTTTCATATCAACGAATGTATCGTAGTCATTCCAATCAGGA
 GTGATAGCAAAGTTGTATTTAGGTGGAATAGCTGATGCTGTAGACCATTCT
 AAAGTACGACCAAGTCCATCCAGTTATCTCCAGTAGCTTCACGTGGAGCT
 50 TTGATATGACTATAAACGATACTTGCAACTAGGAATAAGAATCCAATTGCC
 ATCAATACTGACCCGATAGTTGAGATGAAGTTTAGTAACCACCAACCATCA
 GAAGGCATGTAAGTGTATAGACGACGTGGCATACCATCTAAACCTAGAATG
 AATTGTGGTAAGAAACAAACGTTAAATCCGATCATGAAGAACCAGAAGCAC
 CATTTGTTTAAATGTTTCAATTTAACTTGTAGCCCATCATTTTTTGGATACCAG
 55 AAGATTAAACCAGCTAAGCAGGCAAATACTACACCAGTAACCAATGTATAG
 TGGAAGTGAGCTACTAAGAAATAAGTGTTGTGATATTGATAGTCAGCTGAT
 GCCATTGCAAGCATTACACCAGTAACCCCTCCTAATAAGAAGTTAGGGATG
 AATGCTAATGAGAATAGCATAGGTGACTCAAATGTAATTCTACCTTTGTAT
 AATGTGAGCAACCAGTTAAATAGTTTAACTCCGGTTGGAACACCGATTAAAC

ATTGTTGAGATAGAGAAGAATGAGTTAATTAACGCACCATTACCCATAGTG
AAGAAATGGTGAACCCAACTAAGAACTTAAGAAATGCGATACCTGCAGTT
GCCCCAATCATACTTTGATGACCGAATAAACGTTTACGGGCAAAGTAGGG
ATGATTTCTGAGTACATACCGAATGCTGGCAAAATAACGATATAAACTTCA
5 GGGTGCCCCATACCCAGAAGAAGTTTGCCCAAAGCATTGGCATACCGCCA
TTTGCTACAGTGAAGAAC TGAGTTCCAAAAATCTATCAGCAGTCATTAAA
GCAAGTGCTACAGTGAACACTGGGAATGCTAATATAACGATTAATGTTGTA
ATGAATGTTGTTACACTGAACATTTGGCATTTCGCATAAACTTCATTGTTGGA
GTTTTACATCTTAGAATCGTAACAAAGAAGTTGATACCAGTCATTAAACGAT
10 CCGATACCAGATATTTGAATTGCAATTAATAATAGTTAACACCTGGACCT
GGACTGAATTACCAGCAAGTGGTGGTAGTTAGTCCAACCAGCAGCTGGT
GATCCACCTACAATAATGATAGGTTGAATAAAATCATAACAGCGAAGAAT
AGCCAGAAACTAACGTTATTTCATTACAGGGAAGGCAACATCGCGTGACCA
AGTTGTAATGGAATAACAACATTCCATAAACCAAGATAAATGGCATAGCC
15 ATAAATATAATCATAATTACGCCGTGCGTAGTAAATCTTCATTATAGTGG
TTTGCTTCCAGAATTTGTTATCTGGAATTGTTAATTGAGTACGTAACATT
AACGCATCGATACCACCACGAACGAACATTAATACGGCAGAAATTAATAC

20

Sequence 3343

step.1000g08.cons.ok

TGTGCTCTAAAGAAGATCCAAATAACATAGTTATTAGACCAACAAATGTCA
CTGTATAAATATAAACATTACTAAGTCCCAATAAAGGTGTAAATCTAAATA
25 GTAAAAAGATTCCAGCCTTTACCATTTGTTGCCGAATGAAGATAAGCACTTA
TGCGTGTAGGTGCTGCCATGGCCCTTTGGTAACCAAAATATGAAACGGAAATT
GTGCAGATTTGGTAAAAGCACCTAATAATAGCATCAAAATCATTGGTATAA
ATAAAGGATGTCGTGAAATTGCATTGCGTTGATTAAAGGATATCAGTAATTG
TGTTTTGTTCCGTGAATGATATATAAAATGATAAATCCTGTTAATAACGCTA
30 GCCCACCAAAACACTGTAATCATGAAAGATTGAATGGCGCCTAATTGACTTT
CACCATTATTGTACCAATAGGATATAAGCAAGAATGAGGAAATACTTGTGA
GTTCCCAAAATACATACATTAAAGATGGTATTATTAGCTATTACAATGCCAA
TCATACTGAACATAAAATAATAGTAAATAGATGAAAAATCTAGGAAGATTGT
CCGTACTGTGGGATAAATATTGCGTAGCATAAAAAATACACCCACACCTA
35 TTAGCGAAATAATTAAGCCGAACATTAAACTTAAACCATCTAATCTTAAAT
CTAAATTAATATCAATTGAAGGCATCCATGGTAATCGAACAGCAATAAACT
TATTTCGAATCAGATCTGGTATTTTCATAATAAAATATGTAGATGTCACAA
TAGGAGCTATTAAATGCAACATAGCCCGCAACCTTTCTTAGTTGACGATGAC
TCAGAGTGAAATAAACTATGAGCATGACTAATAAATTAGTCGCCATAAGAT
40 ATACTAACTCATTTAACCCCTCCTTTTCTTTTACATTTATATTAAGTGT
AATCATTTTATTAAGTGTAATCATTGTATGGCGATTGAAACATATCGTATT
TCAATAAATCTCAATATTTCTTAAGTTATCTGATATCCAAATATTGTTGA
ATGATAGAAATTAATATTCCCGATTGAAAGAGAAGATAATCAAATGTTTTT
CTTAAATTAGTTAACTAACAGGTTTAAATCCTTATTATGGTGCAACACGA
45 TTTATGATTTGATATTGCTAGTTGATTGAAATGAGTTTATTGAGCATACAT
AGATTCAAGCACTGTGCGCTTTTCTTCTGAATTGGCATATCAGAAAAAGGTA
ACTATTCTTATTATATCCAAAAGGGAATAGAACTTTATTTTATATAAGAT
AGAATCCATGAGGCTATAATTTTCAATCATCGATAAACTTAACTAATTC
ATTCTCATATTGGGAGATCTTAATATTTTAATTCATATATAAATAGTAATT
50 TTAGACACTGATCAATCTTTTAAGTTGGTTATTTTAAATATCAGGTTCCAA
TTTTTTCTTTTACAACCTTAACAATCTACTTACCTCATATACCTTCTAAAA
CAGGTTTCATAAATGATAAATGACGTTTGTATCACATCAAAGGCATTT
GCCATATATTTTCAAAAAATGTGTCTTTTCTTCAATCTATTTTCGATATA
TATAGTGAAAGGAGGTTTTATTTATGAATGAAGTCAATCAAGTGACGATTAT
55 CTTTACTCAAATCAAACATAATTCTGAAGGTAAAGTAGTCGAACATAAGCG
TCGTTTTTCTAACATTAATCCGCTTGTTTCTAATGATCAAATCAAATCTTT
TCGTTCCATCATTTGAACGTATCAGTGGCGAAGAGTATGACAAAGTCGAGAT
CGTTAAACAGAACTTTAATTTAAGGAGGATACATCAGTATGTCAAAAC
TTTGGAATCGTATTTAAATCAAATCTTAATAAACCCGTCAAATTACTATT

ACCCGATTTTAAACGCTATTACTACAGAACAATTAATTAAAGAAAGTATGAA
 CCAACTTTTGAAGTTAGACATTTTAAAGATTAGTTTGGGAAAGCCTGTCAA
 AATTTATGCTGCACAACCTTATGATAAAAGTACAACGTATATTTGAAGA
 TAAAAATTAATTCATAAAACATTATTTTGGACGCTTATTTAAAGTACTAA
 5 AATCATAAAAAAGACTTATCGAAGAACAAATTTTAAATTTGTTCAAGTC
 TAAAAAAGAACTTAATTTACATTTGTTCTTTTATTAATATGAAGTGAGATAT
 CAAATTCAAATGATGTCTCGCTTTTAAATTCATAAACTTGTTTAATGT
 GCATTTTATTTAATGATTCATCAAAATAGTATCGGCTGTATGACACCTCG
 GAACATTACAACTTATTTCTATACTTTGCTTATGGCTCAATCATTTTGA
 10 AACATGATTTATGAATGCAGTTTCTATGATAATCATAAATCTATATGAA
 ATATCTGATGATGCTGACTTCATACGCCATAATGCCTGCACCTATGGAA
 TGTATCGTTCCAGCTGGTATAAAGTAAAAATCACCTGGTTGCACTTGTATT
 GTTCTCAAATAATTTTCAAATGTACCTTGCTCCAATTGCTCTTCAAATCA
 TCACGAGATTTTCGCATACGTACCTATAGTAATCTTGCACCTTCATCAGCT
 15 TCAATGATGTACCAACATTCAGATTTTCCATATTGACCTTCTTCGTGTCA
 TATGCATAAGAATCGTCGGGATGTACATGAATAGACAATGGCGCAGCAGCA
 TCTACAATTTTAGCCATTAATGGAAAAATCTTTACTTGGAAAAATCTCCAAAT
 ATTTCTCTATGATGTTTCCATACCTTGATCCAATGTTTGACCAGCAAAATA
 CCATTTTCAATCACACTTTTCCGTGTGGATGTGCTGAAATTCCTCAACAT
 20 TCACCTATGTGATTATTAGGTAGTTGATACCCAAATGACGAAGATTATCA
 CTGCCCCACTTTTATCCAGAAAAATGGGTTTTAAAAATAACGGCATGTTC
 ACACCTCCATTTATTAAGGCTTAAACTTAAATCATTTATCTATGCTTGAATA
 TCACACAATGTTAAATAATAAACTACTTTTGATACAACATAGCCCTTCTT
 TTAAGATATGTTTGAAGCTAACTATCTTCTCACTCGTATTGAGTGTAG
 25 CATATATGATTTATTTCAAACTACTATATGAATATACTATCATTTTATAA
 ACTTTCTATCTGTGTCCATTTGACGTAGGATTCCTTACATATTATATCAAC
 TGCCTAATAAATGAATAGCATCATTTTACTAATCTAATAAAGCGTTTATA
 TTGATTACTATACCATTATTTTCTGAAAAATGACCTTTTAAAAATAACAAATGC
 TTTAAATTTGGATTCAATAGAATCAAGCATTACACTTTTTCAGTTCAATTT
 30 AAAAGTAGAAAAATTTAGTATAATTATACTCATTAATAATTTAACTTTGAA
 TATAAAAAGGGAACAAGATAAGACTTTTGATTTCCTTATCTTGCCCCCTCA
 AATAC

35 Sequence 3344
 step.1001a01.cons.ok
 TCTTCGTTTAAATGATAGTTCCTACTAATGAAATCGATGGGTTTTTATAT
 CTATAAGCTTTTAAATAAAACCTTTAGCAACTTCTTTATTTGTTAGTGTCTATA
 TTGTACCCCTCCCTTAAATTTATCTGAAAAATTTATACTATTATGATAACTTT
 40 TAAATGTGAAGTGTGCAAATTTGATTAAATATACTTCATTTAGGGTATAAA
 ATAATTAATAATGTAAATTTTGATAGGAGTGGGTATAGTGTGAGATTATGT
 TTATGAGTTGATGAAACAACATTCATTGAGTTAGAAAAATTTAAGAATCAACC
 ACTTGGTTCTGAAACGGTAGAAAAATTAGTAGAGGCGGGACAGAGTGCTTC
 TACATCCAGTTATCTTCAAACCTTATCTATTATTGGTGTGAAGATCCAAG
 45 CATTAAAGCGCGTTTAAAGGAAGTGTCAAGGTCAGCCTTATGTTTTAGATAA
 TGGTTATTTATTTGTATTGTTTTAGATTATTATCGTCATCATTTAGTAGA
 TGAAGTTGCGGCGTCAAATATGGAGACATCATATGGTTCTGCAGAAGGACT
 ATTAGTAGGTACAATAGATGTTGCATTAGTTGCGCAAAACATGGCAGTTGC
 TGCCGAAGATATGGGGTATGGAATTTGTTTATTTAGGGTCATTGCGTAATGA
 50 TGTTGCGCGAGTGGGTGAAATTTTAAATTTACCTGATTATACGTTTCCGTT
 ATTTGGTATGGCAGTAGGTGAACCTTCTGATGAAGAAAAATGGGTCACCTAA
 ACCGCGCTTGCCATTTTAAACATATTTTTCATAAAGACAGTATGATGCGAA
 TCAGCATCAACAACGTAAAGAAATGGAAGCATACGACCAAGTAGTGAGTGA
 ATATTATAAAGAACGTACTCACGGTGTGCGTACAGAAAAATGGTCACAACA
 55 AATAGAAACATTTCTAGGACGTAAAACACGTTTAGATATGTTAGATGAATT
 GAAAAAAGCAGGATTTATTCAAAGATAAAAAAGAACCTGAGACAGAAATACA
 TGTCTCAGGCTAGGGTGGGGCGATATTTCAACACGAATGTTGCCCCGCTC
 TTTTATATAATTGTATGTCGTAATGATCAATGACCACCTTTCTTACATTTAA
 TTTGTCAGATAAAATCGCCGTAATCATTTGAGTCAAATGTTTTTTTGTCCCAA

TGATTCTGTTAAGCGTGCGGTACCTGTTCCAGCTAGCATTGAGTCATTAAACG
 TTAAGTGTGTACGACCCATATCGATGAGAGGTTTCGATAGATATCAGTACC
 CCTGCGAGAGCAACTGGTAGATTAAAGTGTAGATAATACGAGTATTGATGCG
 AATGTTGCCCCGCCACCTACACCTGCAACGCCAAATGAACTTATAATAACA
 5 ACAGCAATAAGTGTAAACAACAAATGGAAGTCAATTTCTACATTTGCTACT
 GGTGCCACCATAACTGTCTAGCATAGCAGGATAGATTCTCTGCACAGCCATTT
 TGCCCTATGGATAAAACCAAAAGTTGCAGAGAAGTTTGAATTTCCCTCAGGT
 ACACCTAAACGTTTTTGTGTTGCGTTTGAACATTTAACGGTAATGCACCTGCA
 CTTGAACGTGAAGTAAATGCAAAGATTAGTACTTCTATTGTCTTTTTCACG
 10 TATTTAACGGGATTGATACCTAAGACACTCAGTATAATTAAATGGATAATA
 TACATTGTGATTAGAGCTGCGTATGAAGCAATTAAGAATTTACCTAACGTC
 CAAATTGCAGAAAAATCACTTGTTCGCAAGAGTAGAAGCCATAATAGCTAAA
 ATGCCATAAGGCGTTAATCGTAAAAACAAAGTTACGATAGCCATAACGATA
 GAATAGATTGCTTCTATACCACGTTTAAAGTAAGCTTCCATGTTCCGGCTGT
 15 TTTCTTGCAACTCTAAGATAAGCAAAGCCACAAACGTTGCAAAAAATAACA
 ACTGCAATTTGTGCAAGTTGTACGTTGTCTCTGTGAAATCTAAAAATGGATTG
 CTTGGGAATACTTCGAGAATTTGTTGTGGTAAAGTGTTCGAGTTAAATCT
 TTGGCTTGTTTTGAAATTTCTGTACCACGTGAATGTTCTGCCTACCTAAA
 TCAATAGCATGCATCTAAACCAAAGATCAAAGCGTAAAAAATTCACACG
 20 ATAGCTGCAATGGCTACAGTACCAATTAACCAATAAAAAATATAAGAACCG
 ATCTTAGCGAATTTTTTACCAATTTGTATTTTGTAAAAGCGGCAACAATT
 GAAATGAATATTAGTGGCATGACAATCATTGTAAATAGTGAACATAACCA
 TCTCCAACAATACTAAACCAGTCTGTTGATTGTTTCGAGAGTTTTCAGACTCT
 GCACCATATATGAGGTGTAATACAATGCCAAATACGATACCTATTGCTAAA
 25 CGCGTAAAGACACGTTTTAGGGAATGAGACGTGTTTTCTTGCCATAATATTT
 AACATTATTAGGAAAAACAAGTAACACGATGACATTAATTACTGTAAAAAGT
 ATTTCCATATCGATATATAGCTCCTTTTTTAATTTTATGATTCCGACATAT
 GTAATAAGAATAAAGGGGGTTAAGAAGAAATTCAGTTGTTGATTATATAGT
 TTTTCAGATTCAATTTAAAGTTGAATAGAGTGAAGGTGATTACTAGAACAATG
 30 AAAAATGAGAAGTTTATGTTTATCAGAAAAGTGAGCAAAAGAGGACGTAAAG
 AGGTTTATATTAATAAAAAAGACGCGTGCAATCATGCTATAAGAAGACATG
 ATTGCACGCGTTTTTATTATTAATATTATGAATAAACATGAGTCAGCTATGAA
 TACAAATCATTTATAAGTTAAATAGACGTGCAAGAAACCTTTCTCTCTTT
 CTTTAGGTTGTTGTTGGTAAGTCGATTGAGAAGAAGGTGTAGCTGATGTTG
 35 ATTGAGAAATGAGAAATGACTCAGATTTTTCTGAGTTAGATACACTTCCTG
 ATTCTGACAAGAACTCGCTTGTAGACGTACTTTTGTGATTCAATTTCCGATT
 TGCTATTTGTATTACTTTCCGGATGTTATTGAGTTAGACTCGGACAGAGATG
 AACTCATTGAATCATTTGCTGATTGAGATTGGCTCAAACCTGCTGTAGTAC
 TTTGACTTGCGACGTTGCTCTGAGATATTGAAGCGCTATTGTGACATTGAAT
 40 CAAAATGACTACGTAATTCAAAATATTACATTTAATAAAGATATAACACCT
 AATTACGTTTGTACGTTAATAAATGAAGATTCCAGAAAAATAATTTAAGT
 AATGTAGACAATCCGAATATCATTATTGAAATGATTCTTATTAAATCAAAA
 AAATTGCTTTTTTAATGCTATAAGTAAAGAAGGCTTAGGTACCTCGCCTAAA
 ATCACATTTACAAAATAATTTATACCCTGTAACCTCTTTTCAATAAAAAACA
 45 ATATGATATAACTAAAAAGCTCCCAGTAAAGGTCACATTTTAAATGCT

Sequence 3345

step.1001a04.cons.ok

50 ATAAGCTGTTGTGGCAGCAGCAAAATGTGCATCCTGCACCATGATTATAAAT
 TTGTTGGAACATGTCTAGTGTAAATGATAAAATTTGTTGGCCATCATAGTA
 CAAGTCATAAGATTTATCTTGATCGAGTGCCTTTACCACCTTTAATAATGAC
 ATGAGGTGTGCCTTTGTCTATAAATCACTTGAGCAGCTTTTTTTCATATCCTC
 AATTGATGTTAATTTTCTAAACAGAGAGTTGACCTGCTTCGAATAAAAT
 55 CGGGGTAACAACTGTAGCTTTAGGTAGTAAATATTGAATCATTGCTTCTGT
 GTTTCTGGGTTAAGTACTTCGTCTTCTCTTTTACAAACCATTTACTGGATC
 AACTACAAAATAGTCTGCACACAGATTCAACAAAAACATCTCCGGCAGCTTT
 AATAATGTCTTGTGTCCCTAACATTCCTGTTTAAATAGCATCAGGTCCAAT
 TGATATTGCAGTTTCAAGTTGTTTTTCGAAAACATTATCAATAGGTGT

TACATCATGGGACCATGTTTCCTTATCCATTGTTACAATAGATGTTAAAGC
 GACCATTCATATACATCAAGTTCCTGGAACGTTTTAAGATCTGCTTGCAT
 ACCGGCACCAGCACTTGTATCTGAACCAGCTATCGTTAATACTTTTTTTAA
 AGCCATCATTCACTCCCATTAATTTCTAGTGTCTTTATCATATCATG
 5 TTTATCGCGTACGCTAAATTATTATAATTTTAAAATGCAAATCAATCATCA
 TACTTAGCTATTGTATTGATGTTATCTTTAATTCGTGATTAATAGGTAAGA
 ATGTCATCTCATTCTTACTTAAAAGGTCTTCAGTTAATTTTATTTTGAAAA
 TTTGATTGCTAAAGTAAGGCTTAATATAATAACTTCTTTCTTTACTTTCCA
 TCACTAATTGATATTGTGTGTAATGTATTTTATTGGCATCGATAACTGCAC
 10 CTTGAGGGGATACTGACTGATTCTAGAACTTTAAAACAATTCATTAATTTT
 CATCTTCATTATGGGAACAGCGTAGTTGGTGTCTTAAATATGTAGCGCGTA
 TAAAACCCAAACAATAAACTCCTCAGGTGCTAAGTTAAAATAACCATAAC
 CACAGCCCATGTATCTTACTAGCACTTTACCTATTAATTGATTGGTTGATT
 TCTGTGGCGTTAAAAAAGCGTAATTTCTTAAGTTAGATAGATGCCAATCTA
 15 ATTTAGGTTCAATTTGTTAAGGTATGAACATAATTATCTTTAACTATTAATA
 AGCCATTGTGAGGTCTATGGCTACGGTATGTCTGTTTCATCAGTGACCA
 TGAAATGTAAAGGAGGAACGATATTCAAAGTCGTATTTTTTCATTCATAA
 TATTGATTTTCTTAACCTTTTGTCTTAAATTCGCTAATACTTTTATTTAAAC
 CTTAAACCCAAACAATAAACTCCTCAGGTGCTAAGTTAAAATAACCATAAC
 20 GTTTATGGGTACTGTATGAGGCTTCACCAGTGAAGTAATGGTTCGAAATAG
 CTAAACCTTTTTCGTTTATACCATCACCAATCTATAACGTCCTACTTTTA
 AATTTGTTCACCAAAAACCATATTCAAGACGCATGTCTGAATCTAGATCAA
 ATGGTAGTGATAATGGCGTGGAAACAATGGTTGGGATACCATTAAATTCAA
 ATGCAAAGTCCATTGTTCTAGCTAAATAATGGTAACGTTGTTTGTATATA
 25 AAGAAATGGCAGTACACATATATAACACTCCAACTACTGAAATTTATCTA
 TATTAATTTGTAATATAGATAAAAAATGCGCTTATTGTAAACCTGAATCATAT
 AAAAGTTTACAAAAATTAATGCTTATAAAATCAGTCATTTTCTAATATGA
 ATATGTTAAATGTAGCTTGAGGTGAAAACGAATGAAATGGTCAGAGGTAT
 TTCATGATATAACAACGCGCCATGATTTTCAGGCGATGCATGACTTTTGTAG
 30 AAAAAGATATACGACTCAAACCGTCTATCCAGATATACAAAATATCTATC
 AAGCATTTGATTTAACGCGCTTTGAAGATATCAAGGTGTTTATTTTAGGGC
 AAGATCCTTATCACGGTCCATAATCAAGCACATGGTTTAGCATTTTTCAGTGC
 AACCTCATGCTAAATTTCCACCATCTTTAAGAAATATGTATCAAGAACTAG
 AAAATGATATAGGGTGTATAGAACTTCGCCTCATTTACAAGACTGGGCAA
 35 GAGAAGGTGTCTTGTATTAAATACGGTATTGACTGTTTCGACAAGGTGAAG
 CACATTACATCGAAATATTGGATGGGAAACATTACGGATGAAATCATAC
 AAGCTGTTTCTAATTATCGTGAGCATGTTGTTTTTATTCTGTGGGGAAGAC
 CGGCTCAACAAAAGGAACGATTCAATGATACATCTAAACACTTAATCATTA
 AATCGCCACATCCTAGTCCACTATCGGCTTTTAGAGGATTTTTTGGTTCTA
 40 AACCTTATTCAACTACAAATAACTATTTAAAATCTAAAGGGAAAACACCAG
 TTCAGTGGTGTGAAAGTTAGGTGAGAATAATGAACAAAGAACAGATTCTAC
 AATTGATTGAGCAAGAATTGATACAAGCAGATGAAGCTCAGACAGATACGG
 AATTTGAAAAGCATATGTATGCTATACACATGCTCACATCTCTTGTAGTT
 CTCATCAAAGTCGTTCTACAATAGAGAAATTAATCATTCTAAACCAATGA
 45 ATAGTAATATCAAAGATGATTATGAGATGAAACAACAGTCTTCACAAAAAC
 ATCATGTAAGTGCAGCTGAAATAGAAGCAATGGGTGTTAAAGTACCACAAT
 CAATGAAAAAGCATCATACTTCTAATAATATGATGATTACAGATGATCAAG
 TTGGTAATGGTGAATCTATTTTGTATTTTAAATATAAAGGAGTACATTCA
 50 AAATGATGAAAGTTTTT

Sequence 3346

step.1001a09.cons.ok

CTTCTTAAATTACGTAATTTGTATCCAATATCCATCTTTCTAACCACCTAA
 55 AAAAGCTATATCATGTTTACTTATATCAAACATTTTGTTTAATGCTTAATA
 AAAATAACAGAATATAGCTACGTAATCAATAGTTTATATTAACTTATTAT
 TGTGGATTTAATAGAAAAAGTTAGTCGAAATCTTTACTAAATAAAGTTTGA
 TTTGGATTTTTTTAATTTCTTGAACACTTTGTAACTATTATAACCACTT
 TTATTTTCAAATCTTTAAAAATCTGTTTTTCATCAGCTTTTCCAGGCACA

ATATGTTTAAAAATTTTATAAGCATTAAAACTTCGTACACCTTTTACAGAT
GCCTCATAATAATTTTCGATTTTATTAAAGAAATGAATTACTTTTCATCATT
TCATCATTAGTCCAATCTAAATCTAATGGATATTGATATTCCATAGTGTCC
TCTCCTTCATATACTTTTATTTTTCACATTTAGATATATTAATAAACAC
5 AGTATACCTTTATTGGCATAAAAAGATGGGATAAAATCCATTTTGAATTC
TATCCCACCATGAATAAATACCCCTAGCAGATTGCTATTATTTAGAGTAGC
TAAACGCTAAGGGTATTTTGAATGATTTAAGATACGTATGTCTTTGCCAG
TCCGTGACTATACAAACGAAAAAGTAATATACGTACAAATCATAAATAAAT
CATATTAATTTATATACATTTATCATTATTAAGTAGTAATTACATAGTATG
10 AATTGGATAACCAATTGCTTTTTCAGCAGCTTCCATTGTCATTTACCTAA
AGTTGGGTGTGCATGTACAGTTAATGCGATATCTTCAGCATTACATACCTGA
CTCAATAGCTTAAACCTAATTCAGAGATAATATCAGATGCGCCAGTACCTAC
AACTTGTGCTCCAATAAGCGTATCATCTTCTTTAAGTGTAATTAACCTAAC
AAAACCATTTGTATCATCTAATGATAAAGCTCGTCCATTAGCTGCATAAGG
15 GAATTTAGAAGCTTTAATTGATAAACCTTCTTCTTTTGCTTGAGCTTCAGT
ATAACCAACTTGTGCTAATCTCGGTTCTGTAAAGCAAACCTGCTGGCATACC
AATATAGTCTACCTCTCGGCTTGACCATCTATCGCTTCAGCAGCAACTTT
ACCTTCATAAGCTTTTGTGAGCTAATGGTAATCCAGGTACAATATCTCC
AATCGCAAGATATTTTCAATAGAAGTACGACTTTGTTTGTCCACTTCTAG
20 TAATCCACGATCAGCAAAATTCAGACCAAGTTCTTCTAATCCTAATTCATC
AGTATTAGGGCGACGGCCAACTGTAACCTAATACATAATCAGCTTCGATAGT
TTGTTCCCTCACCTTTTGCCCTCATAAGTTACTTTGACACCATTTTCAGTTTC
TTCTGCAGATTTTGCCATTGCTTCAGTAACGATTTCGATACCTTTTCTTT
CATACCTTTTAAACAGGTTGTGTCAATTTGCTTTTCAAATCCGCCATAAAT
25 ATCTTTTGCACCTTCAAGGATAGTAACCTTCAGAGCCAAAGTTTGCAAAAGC
AGTACCTAATTCAGAACCAGATATATCCGCCACCAACTACAACCTAGTTGTT
AGGTACTTCTGTAGATTTAAAGCTCCTGTTGAATCGATAACACGTTTACC
AAATTCAAAATTTGGAATTTCAATTGGTCTTGAACCTGTAGCTATAATCGC
ATGTTTGAAATTTGAAGTTTGAGCACTCTTTTCGTCCATGACACGTAAACT
30 ATTGTTATCAACGAAATAAGCTTCACCTCTACAATCTCTACTTTGTTACC
TTTTAAAAGTCCCTTCAACACCGCCAGTTAATTTATTAACCTACAGAAGTCTT
GAATTCCTTGAACCTTTTGTAGTTTAAACGAAACGCTTTCAGCAATTACCCC
TAAGTTTCTGAATTTTGCCTTCAACAAAGCGATGAGAAGCATGTAGTAA
TGCTTTTGAAGGTATACAACCAACGTTTAAAGCATACACCACCTAAATTACC
35 TTTCTCAACGATTGTTACCTTTTGTCTTAATTGAGCCGCGCGAATGGCTGC
GACATATCCACCTGGACCTGCTCCTATTACAATAGTATCTGTTTCAATTGG
GAAATCTCCAACCTACCATGTTTAACTCCCTCCATTAATAATAATTCTGGAT
TATTTAATAAGCGTTTAAATGTGATTATAGCATTTTGTCTCAGTAGCACCAT
CGATTTGTCTATGGTCAAAGCTTAATGATAAAGCTAACACTGGTGCAGCTA
40 CAATTTCTCCATCTTTAACGATAGGTTTGTGAGCGATACGGCCAATTCTTA
AGATAGCTACTTCTGGGTGATTGATAACTGGAGTGAACCATTTGTCACCAG
CGGAACCGATATTACTAATTTGTGCATGTTGCACCTTTTCATTTCTGAAG
TTAATTTACCATCACGTGCTTTTACAGCTAGTTTCATTAATTTTCATCAGAAA
TTTCGAATATTGATTTACGATCGGCATGTTTAACTACTGGTACTAATAATC
45 CTTTATCCGTATCTGCAGCAATACCAATATTCCAGTAATGTTTGTGTACAA
CCTCTCCAGCTTCTTCATTGAAAGAAGTATTAAGTGCTGGATATTTTAA
GTGCAGAACTAATGCTTTAAACATATGGTAAGAAAGTAAGTTTGTAC
CTTGTTACAGCAGCAATTTCTTTAAATTTCTTACGGTGATCCCATAAATCTT
GCACATCAATTTTCATCCATTAATGTAACATGAGGTGCAGTGTGTTAGAAT
50 TAACCATTTGCTTTAGCAATTTGCTTTGCGCATTCAGGTATTTTCTGTAG
TTTCAGGAAGTCGCTTCTGGTAATGCTTGTGTTGCAGAAGCATTAACGA
CATCACTAGAAGTTGATTGATGATGCTGCGCTAGTGTGTAACCTTCTTCGG
AACTACCACCATTTAAGTATGCATCGATGCTCTTCTTTGTGATTCGTCCAT
TTTACCAGAACCATTACAGCTTTAATATTGACACCATTTTCACGTGCAT
55 ACTTACGCACTGACGGCATCGCTTTAACAGTTTACTTTTCATCTACTTCTG
TCTTTTCTTGTGATTGAGTTGATGAAGCTTCTTCTGCACTGGTGATTCTT
GTTCTTTTCTTCTTTCTTGAATCCTCATCATCGCCATGACCTTTAAAT
GCATTTCTTCTGCATCAGGTGCATCAATTTAACGATGACATCTCCTACTA
CTGCCACTGTTCTTCTACTAACAACCTTCTAACAGTACCACCTACTG

GAGAAGGAATTTCTACTACAGATTTATCATTTTGAACCTCTGCTAATACAT
CATCTTCTTCAATTGTATCGCCGGCTTAAATAAACCATTTAAACAATTCAC
CTTCGTGGATACCTTCCCCGATATCGGGTAATCTAAATTCAAATGCCACGT
TCTTGTCCCTCCTAAGATTTACGTATTTAAATTT

5

Sequence 3347

step.1001b03.cons.ok

10 AATTTTAACTCTCAACAAAAAGTCGTGATGGAACAATTTTTTACACCTAC
ACAAAAAGTTAAATTTCAAAAAATATTACAACCTGAGCACGAACATCCTAC
GGTACAATCTATCATTTTATAACTTTCACGAGACGTTTCGTTTTTCAACGA
TGAAAAATGAACTTTTAGCGTTTGCAATTAATGCGCTATATCATTTAGGAGA
CGTATTTTTTATGTGATAAAAAACATCGTTACAGGGCCAAATCATTGATCAAAC
15 TGATACTAAAAATACCAGTTCTCGCCGTTTTCCACAGTACCCACGTAAAAAA
TATCAATGATATATATCATTTCTGAAATCAAACAAGCGTATAAACCTGTTTT
AGATAACTTATCCCGATATTTCAGGAATCATAGTATCTACTGAACAACAAAA
AACGATTTTATCTGTAAAAATTAATAACGTTATTTCCCATTTACGTTATACC
TGCAGGTTATATTGATACAAATGAATCTCATCATAGTAGTGACAATAAACC
ATTGCCTAACAAAAATGATTTCTATCGGCCGTTATTTCTCCTGAAAAGCAATT
20 AGATCATCAAATAGAACTAATGTCTAAGCTAGTTCCAGCATTTCCAAATTT
ACAGTTACATTTATTTGGGTTTGGTAAAGAAGAAACACATTATCGTAAATTT
AATAGCTCAATATCATTTAGAAAATCACGTGTTTTTACGTGGCTTCATTTA
TGATTTAAATCAAGAAATAGAGACCGCCTATTTATCTTTATTGACAAGTAA
AATGGAAGGGTTCAATTTAGGTGTACTTTGAAACGATTGCTAAAGGCGTACC
25 TACAGTAAGTTATGATACCAAATACGGACCTTCTGAGTTAATTGTTAATCA
TAAAAATGGCTTTTTTAATTGAACAAGATAATAAAGAACAACCTCTATCACAG
CGTTAAAAAGTTATTAATCTGATTTCTAACTTAAGAGAACAATTTTCTAAGGA
AAGTATTAACATGCCCCAAATATTTAATGACAAAAATGTTTTTGATACATG
GCTCACTGTTTTTCAGAACGTTAAAGTTAATTTATAATCGCCAATTTATTA
30 AGATATATAAGGGTGAAGAGAAGTGTGAAACATCATACATCATCTTCACCT
TTAATTTTATTTCTATAATTTTTCATCTTTAATCTGTGCCTTTTGATATTGAA
CATAGCCTGATATATAATGTTTCGATATCATCAATCCGAACCTCGATACCCAT
GGTGTTTTTATGAAGAACTACCAAGAATTCGCTTTGGATTTTTGAAGTACA
TTGCAATTTTCAGGATAGAAAAATCCTACTCGTTGATATTTCTGCTCTTGAT
35 GAATCACATTGATTAATTTTTCTTCATCAATCAACTGAGTCACAAGCTCTT
GACGATGTGTAGCTTTTGCTTTTTGAATAAGCTTATAGGTTGCCATTAAAC
TTCTATAAAAATGTTGGGAATGTTGTTTCACGTTGTTCAATATATTTCCAAAT
AGGTGTTTCACATTCTTAATCCCAAATTCAAAATATTTATCTTGCGGACACA
ATTGAACCTAAGTTCGTTGCAATACCCAAGCCAATGATCATGGTATTGCC
40 AATACTCTTTTTCAATAAATCGATCCATTAACCTTCTCACAACCTTCCAAAC
ATTTATCATTTATGATCTTGGTGATATAAGCGTAATAATGCTAAAGCTGCTT
CACCATCATAATAAATAAATTTCTAAATGATTCTTTCACAGTTAAATCCGGAT
AATTTAAAAATATGAGTTGTTTCGTATGTATCTTGATTAAATCATTGATAAAA
TTCCTCGAGCAACTTTTTGTGCCACGCATAAGTATTTGCTTATTGGGGTTAT
45 GCTTTAAATATTCCAAACCGCAAATATAAAGGCAGCATTTTGTCTTAATTT
TTATTTTCGTTAATATCTTTTGTATCATCAAAGATATATCCAACACCTTCAT
TATCATAGAAATAATTCTCAATAACGTAGTTAATTGCCTTTTCGACTATAG
TTAAATCTTCTCTAAATAATCTAAACCTCTATTAATGCATAAGTTGAAG
AAGCATGCTCTTAATATATTATAGAAATGATTCTTTATCAAAATGTGGAA
50 AATAACCATATTGATATCTTCCAGTATCTGATAGCATATTTCCTAGGAAAT
ATGTACCATTTTCAATTAATTTGGTCTATTTCTTTTATGTAAATAATCGACCT
TTCTTAATCCTTTTTTATAACCTTCATCGTGTAATTCATATATCTTTTGT
CGTCTAAGATAAAAACCTTTAGTTTTGAACTTAATGACTTCTTTGTTTTCAT
AAAAATCATAAGCAAACCTTTTTCTTATGGTTCGTATACCTTTCTTAAATAGT
55 TATTTATATTTTGTTCAGATAGAATAAGCTTCGTTTTCCCATCTGTTTTCA
CTGGTTTTAATAAATGCATTAGTGTTTATTTCTTCAGGTAAAAATGATAAAT
TCCAGTATTGATCTAAAGCTATACCAAAATCAATATAATTTCTTCTAGTTT
GCGTTAATTCATCTTTAACATCTTTAAATAAAGTAACCTTCTTCTTCAGTAA
CAATGTCTATTTTTACCCATGAAGGAAACGCTCCTGTTTTCTGTCTAAACT

5 TTTGAATTTCTTTAGTAATATCCCTTCTAAGGTAATTAGGCTTTTTTAATA
ATTTAACTTTGGCTTTTATATTAGATTTTCCAAGGCTTAAAAAGCATATT
CTTTTATTTCTATTGTTTTTCAACATGATTAATCAATCTTTCTATATGAT
TTTCCATACTCATCTTTTTTCCTTTCAATTTAATATGATTTCAAACCAGG
CCATATAGACCTGGCTTGAATAAATTATGAATGAATCAATAGCGCTATTT
TTCTTTATTATCTTTTTTACGACGCTTCTTAATAATAATGCTCTAAACC
TGCAAATAAAGTTCCAAGTAATGTGCCTTTAGAATCATGATCTTCATTTGC
TCC

10

Sequence 3348

step.1001b04.cons.ok

15

CGTGACTATATGATGGTAGAAATGTGTCTCGTTGTAAAGCGCATTTTAAC
AGTCGTTGTAAATTTTCAATTATCATTTATATTTTGAAGTATAAACTTGT
TACAAAGTTTGTGAATATTTGAGCTCCTAAAGTGAATTTTATTTAAAAAT
TCGCTTTAGGAGCTCTTCTTATATCAACATTCTATTCATGCATTTAAAGTT
TTCTTCTTATACCTCGGTTGATTTTCTCTTATTATAAGATTATTTTCAAC
TTTCTCTACTACTTTCATTTAATTCATCAATTTGTTGATAGTTGTCGTTGA
AGATCCTGAAGAGAAATACCATAGTTTTGGATCTAATTCGTAAATATGATT
20 ACTTTTTACTGCTTTTACATTTTTTATAACTTTGTTTTTAAACTTGATT
TGTTGTTGCTTTTACCACCTACAACTGAACCACGATCCATAGCTAAAATAAC
ATCTGGATTCTGCTTGTAAATATATTCATTATTTATATTTTGACCATGCGG
GCTTTTGTCAACCTTTTGTCTGCAGGTTTAAATCCTAATGTATCAAACAC
TAAACCACCAAATCTTCCCTCCGTGTCCAAACGTTGATAGTTCACTTCGTT
25 AACCAATAAATACATTACTTCTTATTAAAGTCTTTAGTTTTATCTTTTCAT
ATCAGATATTTTTCTATCTAAATCTTTATTAATTTTTTAGCTTTATCTTC
TTTATCGTAGATTTTCCCTAAATTTCTGTATTTTTTTTCATATCTTTAAT
TAAGTTGTCATCACTTGTACCTACATATACAACCTTAGCTTTTGGTGCAGC
TTTTTTAAATTCATCTAAATTTTCTGATTAGCTGTTCTTCCTGAAATAAA
30 AATCACATCTGGTTTAGCTGATGCAACTTTATCAAAGTTCACTTCTTTTAA
ATTTCCAGTATTAATATACCTTATCATCTTTAAATTCATCTAAAAATTTAGG
TAAAGATTGGTTATTTTCACTTTAGGTAAACCTTTTACTTTATCAGCCAC
ACCTAATCTTTCAACACATCAAGCGCTCCATAATCTAATACAACGGCATT
TTTAGGATTTCTTTGGTACTTCGACAGTATTAGAGATTTTTTTCTTATCACT
35 GCCATTATTTCTTTACCCTTGCTTCAAAACTATTTTTGATGGTTACAGT
TTCTTTAGAATCACTATTTTTCTTTTTCGAAGTTGAATTATTATTCGAAC
ATTACTACAAGCCGTTAAACTAAACTAGAGACAATAATAAAATAAGAC
TGTTTTTTTTCATATGTGCAACTCCCTAATCAAATTTATAATTTCTTTTCGT
ATGGATAACCTCTTTTACTATGCGTCTCAATTTTAATCATTAAGTGAAGTC
40 AAAAGTAGTTTCATCATAATATAGACAAATACGTTGTCCCTTATCTCCTC
TATACGTACTTCCATTTTCATATAAACTTTTTAAATGTGAGCTGAATAAC
ATTATCTTTATCATCAGCTTTAACTAGCTCACCTTGTTAAGCGCAATGAT
GTCGCTGAATAACAAGAAGCAAAGTTGATATCATGTAATACGATAATGAT
TGTTTTATTTAACTGACGACATAAATCTCGTAAAGTTTGCATAATCTGTAC
45 TGAATGTTTCATATCTAAATTTATTTAAAGGCTCGTCTAATAATATATAATC
AGTATCTTGTGCTATAGTCATAGCAATATATGCACGTTGTCGCTGACCTCC
AGATAGTGTTTTAAATATTACGATGTTTAAATTTTCAATTAAGCTGTAATAAGTC
AAGAGCATATTCCACTTTGTGTCATAGTCTTCTTGTTCATTTCGACCTTTTGA
ATATGGAAAACGACCAAATTTAACTAGTTGCTCGATAGTTATGTTTCATATC
50 AGTGTGATTAGATTGTTTTAAAACTTAAATTTTTAGCTAAATCGTCAGT
TTTATATTCTGAAATTTCTTGTCTTCAATCAAATTTGAACCGTTCTCATA
ACTAAATAAGCGACTGATAGCAGATAATAATGTACTCTTACCAGCACCGTT
AGGACCTATGAGTGAAGTCAATTTTCCCTTTTTAATATCTACGCTAATATC
TTTTAAAAAGGTTTATTGTCATATCGTTTTTATCTAAATTTTGAATTTGAAT
55 CAATTTGCATTTCTCCTTTTGACAAGCAAATATATGAAGTAACTTCCACCA
ACTAAATCTACTATCAAACGAATTCAGTTGTTGCTTTCGAATAGATTTTCA
ACTACCCATTGCGCTATAAATAAACTAATCCAGCTGAATAATATGTTGCC
GGTAAATAAATTTGTGTTTCATAGGCTTTCATAAATTCGTGTGCTAAGTTG
ACAGTTAATAAGCCAAGAAATGTTACAGGACCTATTAATGCAGTAGATATA

GAGACGAGTAAAGCAACCAAAATAAGAAACATACGCGTCATATTCTCATAA
 GAAACACCTAGATTAATTGCTTGCTCTACCTAACAAAAGAATCTAAA
 TATGGTCGAAGTATTATTGTAACAATTATTAATATAACTAATAATATGCCT
 GAAACAGTTACTAATTTAGAAATTTGATGCTTCAAAGCTAGCAAACATAACA
 5 TTTTGAACGCTAAAAAGATTTCAGGATTCATTATAAGTTGGAGAAAACCTT
 GTAATACTACGGAAGAAAGTTCCATAATTACACCAACTAGTAATATGAAA
 TAGACTGAAAAGTGACCTAGTTTAAATATAACTTGAAATAATAGCAATGAA
 AAAAAACCACTTGCAGTCAAAGTGATTAATAAATTAAGGTAGATATTGTGTA
 ACAACAGTTGCTTGTTCACCTAAAAATAAAATAGGTAAGACTTTGACAAAT
 10 AAATACACAGAATCAAGTCCCATAAATAGAAGGAGTTAATAATCTATTGTGTA
 GTAATAGATTGGAATAAATACTACGGACGTTCTATCGATGCTCCTACTAAT
 AATATAAGAATAAATTTTTTGAACGGCTTTGGAATTGATATTCAAATATA
 TCAAAATCTAAACCTACCAATAAAATAAAAAATTGCCATACAAATTGCCATA
 ACGACCAGTATACTTATTTTTTTATTAGCGCTCATTTCGATAATTTTTTCCT
 15 ACCTTTCATTAGCAAGATTAGGAAAAATAATTGTGCCAAATAC

Sequence 3349

step.1001b10.cons.ok

20 CAATGGACTTTTATCGACTTGAGTTTCAGAAGCAAATATAAAAAAGAACC
 TCGAGCCATACCTATTGTTCCATAAAATAAAGAAATAACATTAATATA
 TTCCATTTTCAGTCATACGGTGTATTTTCTTCATGTTTCTACCCACTTTCTA
 TCATTAAGTGCAGCACAAATAGCATTAGCATAATAATACATATCTTCTGTC
 CAATTATGTGTATTTTCTCTATTTCTTTTACTAAAAATCGTTTCGACTAAAA
 25 AATTCTTTTCTTTGTTGAGTCGATATTTCAATCTGTACACCCATGCCATAG
 GCATTTTTTGTGTTATATTATTAATTTCTCTGCCCGCAATTCTGTCAGGA
 GCAGCTTCCACATTAACACCTGAAACTTTTAAATTTGTGAGTAATAAGAGAT
 ATAAGTCTTTTCATCCAATCCACCAATATAACTATTTGCTTGATTGCTCGAA
 TAACCATGTACGGCTATCGTTACATTCATAAATTGATFCCAATATAATAAA
 30 TTGGGGTTATCATAAATTTGTTGAAGTGACGTGTAGAGTTCTATTGTTTTC
 GGTTTTAAACCTTTAAAGTAAATAGTTTGCATTTCGATAATTCTGCAACC
 AATAACGCTAATTCAGAAGTGCCACACTCTATACCGCCTCCGTGTATAGCA
 GTTATAAGTGATTTACTATTTCTATCTTGTGTCTCAATCATCCAATCTTTT
 TCATTCCCTCACAAGTTCAGTCATAGATTTATAAGTATCCATCACTTTATTC
 35 ACCTCCTTTATCTATCATTTCTAATTTCTCCGTAAATATAAGCTTCACGACT
 TACTGACCACTGATCGCTCTGAGAAATATAAACTTTAACATCTCCATTAGG
 TTGAAGTGTAAATTGAGCACCACAAGCTTTAACAGGCACCGATCTTAGGAA
 CGCAGTTTGGCTACTTGTAATTAGTTCACTCGGTAACTTCGCTATAACAGT
 TCCACTTTTAAAGTTGTCAGCGTTAATTTCTTAACATCACTTCCTTATAGTC
 40 TTGATGTTGAATGATTTTATAAGCGCAATTAAACCAATTTTGTCTTCAGC
 TTTATAATGCCATTTTTTTATAGCTCCATTAATCAAATCATATTCAATCCA
 GTCTGAAGATTTCAGGTAACTTTCTTGTATTTTCAGTAATGCTTTTACGATG
 TTCATCAATTTGTGACGTATCCCTCTTGACTTAATATATTTTCAATATTCCT
 AACTTCGATTTCTTAAATCTTTTAAATATTGACTTCTTTTGTGTCAATTTT
 45 GTTGTATATTCTTCATAAAGATTTTGAATTGTATTTTTCAGCCGCGTTAAG
 AGATTCTTCGACGTCTTTTAACTTTTACTACGCTGTGATTAAGTTTGTGTC
 CAATCCTTTTTTGTGTTCAATCTGAATCTTTGTAATTCCTTCATCACTTGC
 ATCTTTCACTTTAACAACGTAATCTTCTAAATTATCTAATTGTTCTTGTAT
 TTCAGTGGCTCGTGCATTAATTTGTCTTTTAAATTCATCAAACATACGGAT
 50 ATATTAAATTTAGTTGCACCATTAATTTTATTATTTATGCGTCACCCAC
 TTCGAATTCAAATTCAGTTAAACTGCAGTTGATGATTTATCGCTATCTAC
 CTGATTTTCGATGGTTAATCGAAATATATATTTGACCTATAACAGTTGTATT
 CGTTGAAGCTTGTGAAGAATTCAAATAGGTATAGTTACTTCAATAACACCAT
 TAAAGGGTCAATAAAACGTACATTTTCTACTACTTGATTAGAACCGTTAGA
 55 GGACTCAAGATATATATACGTTTCAGTGTTTTCTTCACTAATTAATAAAGG
 CTTTTTGTTTTTAGTCACATAAAACCTTAAACCTGCAGTTTATCATCTAA
 GTTATAAAACCGATACCTTCATCAGATATCGGTTTAAAGTAAGGTTCAAT
 TTTAACTTCAATTTTACCAATCTTTTAAATTCATTTATTTTCTCCCA
 TTTATAACATTTCTTAAATCATCATATATAGTATTAGGGTAAATTTGT

GTGAACGTATCTTCTTTAAGATTACCGTATTCTCCAGATTTTAAATATCTGA
 ATGGAATTAGCAGAATGCGTAGGAGTGAAC TTCGCAAATAATTTAATTTGT
 TTGATTGTTACAATACCAGCACCTTTTTCTTAAACATCTAGTACCGGCATT
 ACTGTATCAGTACGTTTTGTACCTGGTTCAGTAATCGTCGACATTTGAACT
 5 CCAGCAGCAGCGTAAATAGGGAAAGTTGTATTTCTTTGTTGAGTCTATGT
 TCAATTGAAAATAAATTACGTTTCTACTTTTATTAATAATCCAAAGAATGGA
 TGATAATTCTGAACAATATTAGGATTAATCCCCACTGTGACATCTCTATCA
 ACGTTTATAGTCACATATCCATAGAGCTCTACAAATCCATTTGCTAAAAC
 TTGAAACGCTGCTGACTCATCATTTATACGTTGCCATTCATTCCTTCTGCA
 10 CGTAGGGTAAAGACTTCAGATGTTTATTGTTAAATCTATCATCATATACC
 ATCACTTTAATGAAAGGATCTAATGATGTTTGACCATTTTGATAAGATTCA
 AATTGTGCCATTCTAAAAATTACATTACCTACCCAGCGTACTGAACGCTCTC
 ATTTCTTCGGGACTTTTCTTTTCCGACACGTTTTCATTATTATTCTGAA
 AGATATGACGTTGTTTTATTTCGGATTCTACCCAATTACTAAAAGATGAT
 15 AATGTACTGGAACCCCAACAACGAGGTCTCCGTTACTTTTTAAATGTTA
 ATAAGATGAGTCATATTATTATTTTTTTGATTGGCCCAACGCGGATAAAAT
 AAACAGTAATCATTATTTACTGAAATAATATCATGCAAATCTAAGTGAGCA
 GTTAAATTATCTATACTTTGAACTAAGCTACGCATATTTTGTG

20

Sequence 3350

step.1001c01.cons.ok

CAAAGTGAGAGGCAACACAGGTACAATGGCTCTTAAATAAAGCATAATGAA
 CCCAGGAATATATCCGAATTGTCGAAATGCGTCAAACCATTTCTTCAACTG
 25 AGAAAGCGACAATACATCAAATCCTTTCTTAGGGTGTATTACTTATTCAT
 TATAGTTGTAATGATATTTTTTGAAACATTAAAAATTAGGTCAACAACATA
 AATCAGACTTAGGTAGGAATTAATAAATCTATATATGATATCTAATTCAT
 CAATTTTATTATTCCATTCTTTATCGAAACTCTAATGAATTGAAAATTTAA
 TATTTCTATCTGAATTTGACTTCTTATTAATCTGACGTTAATCAATTTTC
 30 ATTTACTCTTATTAATACACATAATAAAAAATGAGCCTAGATAATCCTAGGC
 TCATTTTTAAGTTCAACTTAAAAATTATAAACGATTTTCTAATTTCTTTTC
 TTATCTTCAAAGCCTGGTTTACCTAATAGCGCAAACATATTTGTTTATAG
 GCTTCAACTCCAGTTGATTAAATGGAATTAACACCTAATTGATATCCACTC
 ATTGCACAAGCTAATTCAAAGAAATAAACAACATATCCAAATGTTTCTTCA
 35 TCTAACTGAGGAATATTTACAACGATATTTGGAACGCCACCATCGGTATGT
 GCAAGTAATGTACCTTCAAATGCTTTAGTATTCATTCATCGATTGATTG
 CCAGCAAGATAGTTTCAGTCCGTCTAAATCATCTGCATCCTCTTCAATTTTG
 ATATCATGTTTTGGATGGTTGACCTTAACCACTGTCTCGAATAAGAAACGA
 CGGCCCTCTTGAACATATTGTCCTAAGGAATGTAAATCAGTTGTGTAATTC
 40 GCACCTTGATGGATAAATACCTTTGAAATCTTTCCCTTCTGATTACCCGTAT
 AATTGTTTCCACCATTCTGTTGAAATACTGCATAGAGGGTTCGTAATTAATT
 AACATTTTCAGTAGTATAACCTTTGCTGTATAAAATATTTCGAATAGTTGCA
 TATTGATATGCGATATTTTGATCTAAATCATCAGAAGATAACTCTTCACGT
 GCCTTAGCCGCACCAATCATGATTGATTCAATATTGATACCTGCAGTTGCA
 45 ATTGGTAGTAATCCACAGCTGTAAAGAACAGAATAACGACCTCCCACATCA
 TCAGGTACAACAAACGCTCTATAACCCCTCATTGTCTGCTAATTGTTTAAAGT
 GCACCTTTAGATTTATCTGTGCTTGCAAAAATACGTTTCTTAGCTTCATCT
 TTTCCATATTTTCTTCAACCAATTGTTTAAATAATCTAAATGCAACTGCT
 GGTTCGTGCTAGTACCTGATTTTGAAATAACGTTAACTGAAAAATCTTTT
 50 CCTTGTAATAATCAAGTAATCTTTTGTATAACTTGAGGATAAATGATTA
 CCTACAAATACAAATTCAGGGTATTCCGTATTTGTTCTAAATGAAGATGTA
 AGCATCTCGATTGCGAGCAGTGCACCTAAGTATGAACCTCCAATACCGATA
 ACAACAAGTACATCGGAATTTGATTTGATACGTTTAGATGCTTCGACGATT
 CTAGAAAAATTCCTTTATCTATCAATAATCAACAGGTAAATCTAACCAACCTAAA
 55 AAGTCATTACCTGCTCCTGTACCTTTATGAATAGTTTGATGGATAGTTTAA
 ACAATATCCTTTTGTGCTGATCTAGTTTCATGCTTATCAAAAAATCTAAAGTT
 TTGCCATAGTCTAATTGAATGTGAGTCATCTTATTGCCTCCTGTTATAAGT
 ACTTATACTAATTCCTACTAAGTTTAGATGCGCTATTCAATTCTCAAGGCTA
 ACTTATCACAGGAAATATTTTTTATGTAACCTTTTCATTCTAAAAAAGT

TATTTTAAATATCAAAGGTCACACCACCACCTTTAGATATACAATTTACAGA
AATTTTAAATATATATAAAAGCTTTTATAACAGTATTTCAAACAAAAATA
TCTAAATATAATCAGATTTAAAAATGAATATTTATGTAAAATAATGGTTTT
ATTTTCATTATCAATCTGATATGATAGATGCATATTAATAATCATAAAAAA
5 TTATTTGAGGTGT

Sequence 3351

step.1001d08.cons.ok

10 TCCCCCTTCGGGGGACAGAGTGACAGGTGGTGCATGGTTGTCGTCAGCTCGT
GTGCTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCTTAAGCTTAG
TTGCCATCATTAAGTTGGGCACTCTAAGTTGACTGCCGGTGACAAACCGGA
GGAAGGTGGGGATGACGTCAAATCATCATGCCCCCTATGATTTGGGCTACA
CACGTGCTACAATGGACAATACAAAGGGCAGCGAAACCGCGAGGTCAAGCA
15 AATCCCATAAAGTTGTTCTCAGTTCGGATTGTAGTCTGCAACTCGACTATA
TGAAGCTGGAATCGCTAGTAATCGTAGATCAGCATGCTACGGTGAATACGT
TCCCGGGTCTTGTACACACCGCCCGTCACACCACGAGAGTTTGTAAACCCC
GAAGCCGGTGGAGTAACCATTTGGAGCTAGCCGTGGAAGGTGGGACAAATG
ATTGGGGTGAAGTCGTAACAAGGTAGCCGTATCGGAAGGTGCGGCTGGATC
20 ACCTCCTTTCTAAGGATATATTCGGAACATCTTCTACGAAGATGAGGGAAT
AACGTGACATATTGTATTCTAGTTTGAATGTTTATTAACATTCTATTGTAC
ATTGAAAAGTAGATAAGTAAGTAAGATTTTACCAAGCAAAACCGAGTGAAT
AGAGTTTAAATAAGCTTGAATTCATAAATAATCGCTAGTGTTCGAAAGAA
CACTCACAAAGATTAATAACTAGTTTCTAGCTATTTATTTGAATAACAATTC
25 AAAATATGGTGGAAACATAGATTAAAGTTATTAAGGGCGCACGGTGGATGCC
TTGGCACTAGAAGCCGATGAAGGACGTTACTAACGACGATATGCTTTGGGT
AGCTGTAAGTAAGCGTTGATCCAGAGATTTCCGAATGGGGGAACCCAGCAT
GAGTTATGTCATGTTATCGATATGTGAATTTATAGCATGTCAGAAAGCAGA
CCCGGAGAACTGAAACATCTTAGTACCCGGAGGAAGAGAAAGAAAAATCGA
30 TTCCCTGAGTAGCGGCGAGCGAAACGGGAAGAGCCCAACCAACAAAGCTTG
CTTGTGGGGTGTAGGACACTCTATACGGAGTTACAAAAGAACATGTTAG
ACGAATCATCTGGAAAGATGAATCAAAGAAGGTAATAATCCTGTAGTCGAA
AACATATTCTCTCTTGGAGTGGATCCTGAGTACGACGGAGCACGTGAAATTC
CGTCGGAATCTGGGAGGACCATCTCCTAAGGCTAAATACTCTCTAGTGACC
35 GATAGTGAACCAAGTACCGTGAGGAAAGGTGAAAAGTACCCCGGAAGGGGA
GTGAAAGAGAACTTGAAACCGTGTGCTTACAAGTAGTCAGAGCCCGTTAAT
GGGTGATGGCGTGCCCTTTTGTAGAAATGAACCGGCGAGTTACGATCTGATGC
AAGGTTAAGCAAGCAAAATGCGGAGCCGACGAAAGCGAGTCTGAATAGGGC
GTTGAGTATTGTGCTGTAGACCCGAAACAGGTGATCTACCTTTGGTCAGG
40 TTGAAGTTCAGGTAACACTGAATGGAGGACCGAACCGACTTACGTTGAAAA
GTGAGCGGATGAACTGAGGGTAGCGGAGAAATCCAATCGAATTGGAGAT
AGCTGGTTCTCTCCGAAATAGCTTTAGGGCTAGCCTCAAGTGATGATTATT
GGAGGTAGAGCACTGTTTGGACGAGGGGCCCTCTCGGGTTACCGAATTCA
GACAACTCCGAATGCCAATTAATTTAACTTGGGAGTCAGAACATGGGTGA
45 TAAGGTCCGTGTTTCGAAAGGGAACAGCCAGACCACAGCTAAGGTCCCA
AAATATATGTTAAGTGGAAGGATGTGGCGTTGCCAGACAACCTAGGATG
TTGGCTTAGAAGCAGCCATCATTTAAAGAGTGCGTAATAGCTCACTAGTCG
AGTGACACTGCGCCGAAAATGTACCGGGCTAAACATATTACCGAAGCTGT
GGATTGTCTTTGGACAATGGTAGGAGAGCGTTCTAAGGGCGTCGAAGCAT
50 GATCGCAAGGACATGTGGAGCGCTTAGAAGTGAGAATGCCGGTGTGAGTAG
CGAAAGACGGGTGAGAAATCCCGTCCACCGATTGACTAAGGTTTCCAGAGGA
AGGCTCGTCCGCTCTGGGTTAGTCGGGTCTAAGCTGAGGCCGACAGGCGT
AGGCGATGGATAACAGGTTGATATTCTCTGTACCACCTAGTATCGTTTAAAT
CGATGGGGGGACGCAGTAGGATAGGCGAAGCGTGCTGTTGGAGTGCACGTC
55 CAAGCAGTAAGGCTGAGTGTAGGCAAATCCGGCACTCATAAGGCTGAGCT
GTGATGGGGAGAGGAAATGTTTCTCGAGTCGTTGATTTACACTGCCGA
GAAAAGCCTCTAGATAGATAACAGGTGCCCGTACCGCAAACCGACACAGGT
AGTCAAGATGAGAATTCTAAGGTGAGCGAGCGAACTCTCGTTAAGGAACTC
GGCAAAATGACCCCGTAACCTTCGGGAGAAGGGGTGCTCTTTAGGGTTCACG

CCCAGAAGAGCCGAGTGAATAGGCCCAAGCGACTGTTTATCAAAAACACA
 GGTCTCTGCTAAACCGTAAGGTGATGTATAGGGGCTGACGCCTGCCCCGGTG
 CTGGAAGGTAAAGAGGAGTGGTTAGCTTCTGCGAAGCTACGAATCGAAGCC
 CCAGTAAACGGCGGCCGTAACCTATAACGGTCCTAAGGTAGCGAAATTCCTT
 5 GTCGGGTAAGTTCGACCCGCACGAAAGGCGTAACGATTTGGGCACTGTCT
 CAACGAGAGACTCGGTGAAATCATAGTACCTGTGAAGATGCAGGTTACCCG
 CGACAGGACGGAAGACCCCGTGGAGCTTTACTGTAGCCTGATATGAAAT
 TCGGCACAGCTTGTACAGGATAGGTAGGAGCCTTTGAAACGTGAGCGCTAG
 CTTACGTGGAGGCGTTGGTGGGATACTACCCTAGCTGTGTGGCTTTCTAA
 10 CCGCACCACTTATCGTGGTGGGAGACAGTGTACGGCGGGCAGTTTGAAGT
 GGGCGGTTCGCTCCTAAAAGGTAACGGAGGCGCTCAAAGGTTCCCTCAGAA
 TGGTTGGAAATCATTATAGAGTGTAAAGGCATAAGGGAGCTTGAAGTGGCA
 GACCTACAAGTCGAGCAGGTCGAAAGACGGACTTAGTGATCCGGTGGTTC
 CGCATGGAAGGGCCATCGCTCAACGGATAAAAGCTACCCCGGGGATAACAG
 15 GCTTATCTCCCCAAGAGTTCACATCGACGGGAGGTTTGGCACCTCGATG
 TCGGCTCATCGCATCTTGGGCTGTAGTCGGTCCCAAGGGTTGGGCTGTTC
 GCCCATTAAGCGGTACGCGAGCTGGGTTGAGAACGTCGTGAGACAGTTCG
 GTCCCTATCCGTCGTGGGCGTAGGAAATTTGAGAGGAGCTGTCTTAGTAC
 GAGAGGACCGGGATGGACATACCTCTGGTGTACCAGTTGTCTGTGC
 20

Sequence 3352

step.1001e05.cons.ok

TTCATTACCTAATGCAGGTAAAATTTTCTTAATTGCCTGTGGCATAACTAC
 25 TGATTGCATAGTCTGTCTATAATTCAATCCTAAACTTCTAGCAGCTTCAGT
 TTGTCTTTTATCTACAGCATTAAATACCAGCTCTAAATCTCGGCAATGTA
 GGCTGAAGAAATTTATAACAAGCGCAATAGTACCACAAATTAAGCTGAAAT
 ATCTAACCCCAAAGCTGCAGTAGTCCCAAAAAATACAATAAACACTTGCAC
 AAGCATAGGAGTACCTCTTAAAAATTTCTATGTATATACTTGCTATCCATTG
 30 TAATGGTCTAATTTTACTTATTTTAAATAAGCTATAAAGAGCCTAAAT
 AGACCCCAATACCACCTACTAATGAAATCAAAATAGTATTTTTTAATTCC
 TTTAATGAAAAAAGTGCATATTTAGAAATAAAATTTCCATCATCTTGCAT
 ATCTTCGGCTGCTTTTGTCTATGTATTGATCAATTAAATTTTTTCTTTTAC
 ATTATCAATCGTTTGTATTTAGCTTATCTAAAAGTACAGGAGAGTTCTTAGG
 35 AACAGCTATACATGTTTGTCTTTCTTCAATTAATTTAATCTTTGAAAA
 TGTTAGTTCTGAATTTTGTTTTAAATATGCCTCACCTACTGGTTTTTCAAC
 AACTACACCTGCTACTTTTCCACTTTTAAAGATAATATAGCTTCAGGTAA
 TCTATTGAGTGAAGAAATTTTACTATCTTCAATCTCAGTTTGTGCAATTTT
 TTCTTGATCTGTCCCTTTTTCGCGAGCAATTTTTTTCCTTCAAATCTTT
 40 ACCTGTTGTCATCATCATTCACAGGATGTGTGAACATTACTTTTCGTGCTTC
 ATTAAGCATCATCCCCATTCGGCTGTAGGTGCCTTAACACCTAATCCAAG
 GAACGAGAATCTTGACATTTGTAATATCATTTAGCAGATCGAACTACTAGC
 AATAATCGCTATGTCTAGTAAAGGTTAGTGGCAAAATATGTTTGCAGATGAT
 TGGTTAAATCATTTCATACCAATTACTTTGGCAAAATTTTACATGATCAGCTTC
 45 AATATATTGCATTACACTGGTTTGAATCACGCGACAAAACACGCCCCATCG
 AGTCAATATAAACGCAATAATAATATTTTCTACACCCATGCCAAACAACGT
 AATCAATGCCAATGTACACCATAGCTTGGAAAAGCTAACATCACATCGCA
 TATACGCATAATTATTGCATCGATATAACCTGGGAAATAACCTGAAATAAA
 CCCAAGTATCGCTCCTATCACAACGGAAATAATCAATGCGACAAATACATA
 50 TAACAACTAGGTCTTATGGCGTATATTATCCGTGTTAATACGTCCCGGCC
 TAAATGGTCTGTCCCCAACAGTGAGACCAACTTATACCAGCAAAATTTATT
 TGCTGTATCAATGTGATTCGCTTCATAAAATGTAATTAAGGAGCTAAAAC
 TCCAAGCACTACGTAGATTGTAATAATAGCTATAGCAACTATGGCACCTCT
 ATCTTGAAATAGACGACGTAACACAATCATTTTCAATGCTTCCCTTAAACGC
 55 GGATTTAATATTGCATTAAATGATATCTGCAAGTGTATTAAGAAAGAAATTAAC
 AACACCGCGACAATGAGCACGTATGCTTGAATTACTGGAAAATCATGTTCA
 GTAATCGCTTTAACACTTAATGGTCTTAAACCTGGCCACGCAAAATACATTT
 TCAATAACGACTAAACCTCCCATGATCATTTGGTATCGACATACAAAAATA
 GATACGGCTACCTGTAATGCATTACGTAATACATGCAGCAACATTGTCAGA

CGAGGTACACCACTAGCTTTTAAATACAAAACATAATCTTCATTTAATTGT
 TCAAGCATAGATCGTCGGACATTTCTAAAATAAATTCCTGCATAGGTAATC
 GTAATAACAAAGACTGGAAGAATATAACTTTCTGGACCTGTGAGTCCTGAA
 GTTGGTAATAAATTGAGTTTACAGATACATAAACTAATATTGAAGCT
 5 ACCCAATATGAAGGCAGAGCTGTTAGTAGAAAAGCTATTGAACGTATAGCA
 CGATCTGTAAATCGACCACCTAGTCAAAGCACTCACTATGCCTAGAACACA
 GATGTTATCATCACCACCACACTGGAATCAAAGTGAGTTTAAATGTATT
 ATAAACGCTGGACCCATACGCTCAGATACTGGATCCCTGTGATATAACTC
 GTTCCAAATTTAACTGCAAGGCATCAACAAGCCAATTTTATATGCGTG
 10 AATATCGGTTGATTTAATCCATATTTTCTTTAGTTTCAGCTATTAAATCT
 TCTGTTATATTAGGTGTTCTTGGGCATGTAAATTTGTTACTGCTGGATCT
 TCGTTTGTGAGATAGGTCATCATGAAAGTCATAAAGTCACTACAATTAAT
 AATGGAATCATGAGCGTAAGACGTTTCATAATAAATTTAACCATGACGACG
 CTCCTTTATTTATATTGCATTTTCATTAAATGGAAGCTCATATTGTGATTGT
 15 GTAAAGGATAATTTTGTAAATCTTTAGGTGCAACAACAGTCATTCTGCCG
 TGAGTAATTGGAATAAAGACACCTTCATCGTCAACTTGTTTTAAATTTGC
 TGATACGCTTTTGAACGTGCCCTTTCATCTTGAATTTGAAAGCATCATCA
 ATATCTTTGTACAGTTGTGCTTATTCTTAATACCTAATGTTGCACTTTCA
 TATCCCGTTTTCGCTCTAAATGCTGCTAATGTACTCTGAGGATCATATAAT
 20 AGCCCCCATGTTTGATTAAACATTAAATCGTAATCACCAGAACTACGACGT
 TCAGCTACTTTATCTGACGTTTCTCCGTTTATATCTAGTTGTACACCTAAC
 TTCTTAAATCTGCTTCTAAGAATTCAGCTTGTCTTTTGTGACTGGAAGAC
 CCTTTGTCTAATAACAACCTTAAGATTCAAATCTTTGCCCTCTTTTGTACGA
 ACTTGTGCGATCTTTGAAAGCACCCATCCAGCCTTGTCTAATAACGCTTGC
 25 GCTTTTTCTTATCATATGTTCTTGTGGTAAATTAAGTTTATATCTGTC
 ACATTTTGTAGCAATAGTTGTGTGGCTGGCTTTTCTTGTGTTGTCTAAATA
 TCTTGAGCTATTTTATCTCTATTTACCATGTGACCTAATGCTTGTCTGACT
 GCTTTATCACTGACTGCATATCTTTCTTACCAGAATTAACAACAAGCATT
 TTTGTATTCTAGCCTGGCTACGTTTACTTGGTAGCTTCCGGTTTCTTTT
 30 AATTGTTTTAAACTATCTTTATCTAAGCTGTCTGTACCTCTATCATCTGTA
 TAAGCAAAGTTGGTTTCTCCTTTTTCATTGATAAAAAATGTTGTTCTCCT
 GCAGGTTTAACTTTTGCTTCTACTTTATTTAATTTTGCCTTTTCTCCCAA
 TATTGATTATTTTATTAACTCTGCAGATATATCTTTTTTGTGTC

35

Sequence 3353

step.1001e06.cons.ok

TTAAAGTAGGTCAACCAAGAGAAGAAGGAACCTCAAGTAGGTCCATTATTA
 GTAAAAACAATTTGATCAAGTTCAGGATTACATTGATAAAGGTATCAATG
 40 AAGGAGCAGAGTTATTCTACGGTGGTCCAGGTAAACCAGAAGGATTAGATA
 AAGGATATTTGCGCGTCCAACTATATTTATTAATGTAGATAATCACATGA
 CGATTGCACAAGAAGAAATTTTCGGACCTGTCATGTCTGTAATCACTTATA
 ATAATCTTGACGAAGCTATTGAAATTGCTAATGATACTAAATATGGTCTTG
 CTGGCTATGTTATTGGTAAAGATAAAGATACATTGAGACATGTTGCTCGTT
 45 CAATCGAAGCTGGTACAATTGAAATTAATGAAGCTGGTAGAAAACCAGACT
 TACCATTCCGGTGGTTATAAAGAACTGGCTTAGGTGCTGAATGGGGCGATT
 ATGGTATCGAAGAATCTTAGAAGTAAATCAATAGCTGGATATTTTAAAT
 AATTAGTAAATTAATTAATGTTGCAATTAACAAAGCACGGAACATAAT
 TTGTTCTGTGTTTGTGTTTTATGTTAAATTCCTAAAATCATTTTGTAAAT
 50 TAAATGAAAAATTTGCTATAGTTAGTCTAGTAATTAGAAGGTTGAGAAATA
 TACAAACGTATAAATATAGCTATAATGTGACTTTTTATTATTGTCACTCAC
 ATAAATATCGAGATAAGAATCTATTATTTGGTTAAGCTAACTATACGTTTA
 TTATTTTTGAACTTCGTAAAAATAAATAATTAATGAAATGGGTGCAAGCTTA
 TGCCTGGTAGAGAACGAACATCTCCGCAGTATGAATCTTTCCACGAATTAT
 55 ATAAAAACAATACTACTAAAGAGCTCACTCAAAAAGCCAAATCTTTAAAT
 TAATAATTACAGTAAATTAATAAAAAAGAATTGGTACTTGCCATTATGG
 AAGCACAAATGGAAGGATGGTAATTACTATATGGAAGGAATATTAGATG
 ATATTCAACAAGATGGATATGGTTTCTTAAGAACCGTTAACTATTCTAAAG
 GTGAGAAGGATATTTATATTTCTGCAAGCCAAATTCGACGTTTGAATAA

AACGTGGTGATAAAGTAACGGGTAAAGTTCGTAAACCAAAAGATAATGAAA
 AATATTATGGTCTACTTCAAGTTGATTTTGTAAACGACCATAATGCAGAAG
 AAGTCAAAAACGTCCTTCACTTCCAAGCTTTAACACCTCTTTATCCGGAA
 GAAAGAATCCTATTAGAAACGCAATCTACAAATTATTCCACTCGTATTATG
 5 GATTTAGTCACACCAATAGGTCTTGGTCAACGTGGTCTTATAGTTGCACCA
 CCTAAAGCTGGTAAGACAAGTTTATTAAGAAATCGCTAACGCAATAGCG
 AGTAATAAACCGGAAGCGAATTTGTTTATATTACTAGTAGACCTACTCAGG
 AATTAAGAGAGCAAGGTGTACGTGTGAAACTGTCGGCTGTTAAGGATATTG
 TTGATGGTAAAGATATCGTACTTGTAGATGATTTCGATTGTTTCGAGGTACAA
 10 CGATTAAACGCATAGTTAAAAATGCTTAAGGATTTCAGGAGCTAACCGCATT
 ACGTAAGAATTGCTTCTCCCGAATTCATGTTCCCTAGTTTTTATGGTATTG
 ACGTATCTACAACAGCTGAACTCATCTCAGCAAGTAAGTCTCCTGAGGAAA
 TTAATAATCATATTGGTGCAGATTCTCTTGCTTATTAAAGCGTTGATGGCT
 TAATCGACTCTATAGGACTTGATTATGATGCGCCATATCATGGCTTGTGTG
 15 TAGAAAGTTTACAGGTGATTATCCAGCAGGACTTTACGATTATGAGAAAA
 ATTATAAAAAGCATTAAAGTGAACGTCAAAAATCATATATAGCTAATAATA
 AACATTATTTTGTAGTGAGGGAAATTTACATGCTTAAAGCATATGAGGAA
 TCTGGTGTAAATATTCAAGCAGGATATGAAGCAGTCGAAAGAATAACAAGT
 CATGTTGAACGTACATTGCGCAAAGAAGTATTAGGTGGTTTAGGTGGATTT
 20 GGTGCAACATTTGATTTGTCTCAATTAAAAATGAAAGCGCCAGTTCTGGTA
 TCAGGTACTGACGGTGTGGGTACAAAGTTAAAAATTAGCAATTGACTATGGA
 AAGCATGACACAATTGGTATTGATGCTGTGCGCAATGTGTGTAATGATATT
 TTAACAACAGGTGCTGAACCTTTATACTTTTACTATATTTGCCACGAAT
 AAAGTAGTGCCAAGTACTATAGAGCAAATCGTTAAAGGTATAAGTGACGGT
 25 TGCGAACAACCAATACGGCCTTATAGGCGGTGAAACTGCTGAAATGGGA
 GAAATGTATCATGAAGGTGAATATGATATTGCTGGTTTTGAGTAGGAGCG
 GTAGAGAAAGAGGACTATATTGATGGTTCAAATGTTGAAGAAGGACAAGCA
 ATTATTGGTTTAGCTTCAAGTGGTATTCATTCAAATGGCTATAGTCTAGTT
 AGAAAAATGATAAAAAGAATCAGGAGTTCAATTACATGATCAATTTAATGGT
 30 CAAACCTTTTATAGAACCTTCCCTTGCACCAACAAAATTGTATGTAAAGCCT
 ATTCTTGAATTAAAGAAACATATTGATATCAAAGCGATGAGCCATATTACT
 GGTGGAGGTTTCTATGAAAATATTCCGCGTGCCTTCTCAAAGGTTTATCA
 GCAAAAATAGATACACAATCATTTCCCAACGTTGGAAGTCTTTAATTGGCTT
 CAAAAACAGGGCAACATTTCAACGAATGAAATGTATAACATATTTAATATG
 35 GGTATTGGATATACAATTATTGTTGACAAAAAGATGTTCAAACAACATTA
 ACAACGTTACGTGCAATGGATACAACCTGCATATGAAATTGGTGAGATTATA
 AAAGATGATGATACACCTATTTCATTTATTGGAGGTAGAATAGTGACCAATA
 TAGCGATTTTTCCTCAGGATCAGGTAGTAACTTTGAAAATATTGTAAAAC
 ATATCCAAACAGGGCAATTATCTGGTATCAATGTAACGGCACGTATACAG
 40 ATAATGAAGGGGTACCTGTATTGATAGAGCTAAAAATTTAAATATTTCCAA
 TTCATATTAAACAAGCCAAAAGACTTTTCATCCAAATCTTTATATGAACAGC
 ATCTACTTAAATTATTATCCAGTGAGGAAGTTCAGTGGATTGTATTAGCTG
 GTTATATGAGACTAGTTGGCCAAGACTTATTACAAGCTTATGAAGGACGGA
 TATTAATATATACATCCCTCATTTATTGCCTAAATTCAAAGGTTTAGATGCCA
 45 TAGGTCAACGCTTAGAAAGTGGAGATACTGTAACCTGGATCAACTGTCCATT
 ATGTAGATAGTGGGATGGATACGGGAGAAATTATGAACAACAGCAATGTG
 ATATAAACCCGGACGATACTAAAGAACAATTAGAAGATAGAGTAAACATT
 TAGAATATGAACCTTTATCCAGAGTTATAGCTAAATCATTAATAGAGGA
 50 GTATCTTATAGAATGAAAA

Sequence 3354

step.1001e12.cons.ok

TAGTGAATGAAGACCCTGAACGTAAAAATCATTATCGTTTCAGCTCCAGGCA
 55 AAAGGCATAATGACGACATTTAAACTACTGATTTATTAATTCGTCTCTATG
 AAAAAGTACTTAATAAATTAAATTTATGAAAGTAAAAACAAGAAATTATCC
 AAAGATATGCTGATATAGTAGAAGAATTAGGTATAGGAAATGACATTTTAA
 TAACAATTATGACACTTTAGAGGAATACATTAAACATCTTTCTGACAAAC
 CTAACCGTTTATATGATGCTTTATTATCTTGTGGCGAAAATTTAATGCTC

AATTAATAGCCCAGTATAATAATAGTCAAGGTATTCCTACTCGTTATATTT
CTCCTAAGAAGCTGGATTAACTGTAAGTATTTACCACAGCAAGCTCAAA
TTTTAGATTCCGCATATAATGAAATATACAAATTGCGTGATTATGATGAAA
AGCTAATTATTTCCTGGTTTTTTTCGGAGTTTCAAAGCAAAATTATATCGTTA
5 CGTTTTCCACGCGGTGGTTCTGACATAACTGGTGCTATCATAGCACGTGGCG
TCCGAGCCTCACTTTATGAGAACTTCACTGATGTATCAGGAATATATAAAG
CTAATCCGAATATCATAAATAATCCTGAACTCATAGAGGAAATAACTTATA
GAGAAATGCGAGAGCTATCTTATGCAGGATTGGAGTTTTTCACGATGAAG
CTCTACAACCTTTATACAAAGATCGAATTCCTGATGTTATCAAAAATACTA
10 ATCGTCCAAATGATAAAGGGACCTACATTTTACATGACCGTGAATTCGATT
CTAAAAATGTCATTAGTGGAATTAGTTGTGATAAAGGCTTTACTGTGATTA
ATATTAAAAAATATTTAATGAATAGATTAGTTGGATTACACGAAAGATTC
TTGGCGTTTTTAGAAGAATTTAATATATCATTTGACCACATGCC'TTCTGGTA
TTGATAACATAAGTATTATCATGCGTACAAATCAAATTCAGGTAAGAAAA
15 GTCAAGTCTTAATGCCATACGCAAAACGTTGTGAAGTTGATGAATTAAGTA
TCGACCATGATTTAGCAGTACTAATGATTGTTGGTGAAGGTATGAATCAAG
TTGTTGGTAGCATTAGTAAAAATTAATCAGCCCTTTCAGAATCAAACATTA
ATTTAATAATGATTAACCAAGGTGCTTCTGAAATTTCAATGATGTTGGAA
TTCATGAAGCAGATGCTGAAAAAGCAGTATTATCTACGTACGAATTTTGTT
20 ACAACGGTGTGTTGTTTAAAAATTTGTGTAAATAAAATAAAAAAGATTTCT
ACATCACTTTCAATACATAAAGTAAATTTATTAACCTCTGTATCAATGATGT
AGAAATCTTTTAAGATATCTTTTAGTGGAATTTTACTATAGCTTTTT
TAATACGCTGTTACCTACTACCAATTTTAAATATAGAATTGATAATAAC
TAAGAGAATCAATAACAGCTTTATAATTACTAACTGATGATATTTTGTGTT
25 CACTCATTAATCATACATAAGAATAATTTCTGGCTTAATATAATCTACGT
TCCAACTTAAAGAATGAAAAATATATATTATTTTTTGGTAAACGTATATGAT
GATCTAACCTAAATAGCCATTCATCATTTACAACATCATAAACAAATATTG
TCATGACTTCTGTATCGTTTTTAAAGACTTTAACATGTTCAACATCTGATA
AAGATAAGATTCTCTTAATCTTTCATATAAAGTATTATCATAAATAGTGAA
30 TCGTAAATCTTCAGGAATCACCTGATAATTTCCAATAAATTTTTCGCGT
CAACACAAATATCTATTTTCGATGGTAATATAAAATCATCATTCAGAAATA
GATGCATCGCTACCTCGCCGTGAAATTTGAATAAATGAGATGATGTTTTCC
TTAAATTTTCGACTATTTTATTTCACTGGTTTTTCTTTTAAACAAACATGT
TACCACCCCTATAGCAAGCGCTTTCACAATCAAATTTATACATTTTCTTTTT
35 TCAC'TTATCAAACAAATAAATGTGAACATTTAGTAAATTTATTTTATTTT
CAATAAAATATTTTTCATAAAAAGTTACTATATCACTATTCGCTACGCATA
TGATAGGGCTCTCATAAAAATTCACGAAAGTTTAAATGATATGTAATATT
TCGTCACTACTATTTGATAAAACACATTTCCCTTTTCAATTTTCATTGAAC
CAAAATTTAAGAAGTTTATAATTTATCTTGCAAATAGAATTTAATATTTAA
40 AGATTTGGGTGAGTAGTAATGTCTTTAACGATTATCTTATCTATAATAATT
ATTATTTTGATTATGGCTATGGTTCTTAATCAAAATTTATGAAAGATAGG
GTTGAAACAGAAGAATATGCTAGAAATCAATTAATCTCTAAAAATTCATTT
TTAAGTGAAGAAAATTTATCATTTGAAAAACCAATGTTAAGTACAAACAAT
GACGTGCGTCAACACGCTTTTAAAAACGCCAAGCGTGAATTAAGAAAAATA
45 TTAAATAGATTTAAGAAGAGGGTTCGTTTACGATCATATACAATTTGTTCT
ACGAGTAATTTGGCTGTAAACATCCCTTTTTCGAATATGCACGTTCATTC
GATTTTATATCATTTACTGATGTTGGTTTGATAAATGTGGATGTTAAAAAT
TGGAACCAAAAAACGTTTTATCATTTTGATGTGCCAGATCAACATCTTGAA
GAAGGACAACCACAATATAATACCGAAAAAGTTGTCGGTCATTATATTAGC
50 AATCGATATCATAGTCAGTTTAAAAACAACACGTTCTGGTGTCTATACTTTT
ATTGAGATTTTACAGGATAATCGTGTAATATATGAATTTTATGACCACGAT
CCATACGATAAAGCCGCAAAACATGCAAAAGCATTTAAAGATAAAATTGAA
AATGATTATAATTTTAAAAATTCAAAGTATTGGCGTCATATTTTATGATG
GGTAGCGTTAATATTATTGAAGGATCCGACGAGAGTGATAAATACGTCGAC
55 ACCGTATCTACACCGATATCACTTGAAAAAGTAATTGAGGAAGCTATCGAT
TTATCTAAGCACCCCTTACTGATAAAACAAATCGAAGAAATTTCTGAAAAC
TTTAAACAACATATGAATAATTAAAAATTCAAAAAGAGTTTGGGGCATGAAG
TTATTTGACTAATAAAAAAGACAATTTCTATTTATAAAATGTAGAAATTGTC
TTTTTATATAATTATTTAAATTTGTTGAGCAACTACTTTTTTTATATTCAT

5 TAGCATTAATACAAATGCAATTTCTCTTTTGACTTTAATTATGTTTCGAAT
CGACATTCTGATGAACTTAAAAATACTTTTCGTAAATCCATCAACAAGTTC
CACATCAAATTTCTTTAACTGCAAACGTTTCTCTAAAAATAAGAAATCTTA
TCCCAATTTGAATTTTAAAAATATCTTAAATTGTAATTTTACATTCCAACCTT
10 ATTAAATTATCTATTATCAACAATGTTGTCATATATATTTTAAATTTTAAAC
GTTTTAATACGATTTCAGTAGTCATAGCACACATTTATATGACAACACTTAT
AGCAATTATTTTCATGAATAAAATAAACTTAATTATAAACTGTATCAGTTTAC
ATTAGAAATCTTTTTTAAAAACATAATGTTTTTCACTAGACTAAGAAGTAGAT
ATCCTAATTTCTACTTTTATGAAACACTAGTCAGAAACGTAGTTGAAATCT
15 ACTAAATATCACTCTTTTTTTATTCTAATCTCTTTAGCTAGAATAAGAAAA
CGCTGCTAAACAAATGGTCCAACGTAGTTTCTACTTATTGATATAACAAA
TGAATTTTATATTTTCCACTCAATATCAGACTCTAAATTATCGCTACTAGCT
CCAAATATTTAATTTTCTCTTTTTTGCCTCTCTCTGAGCGTTTACGAAGGT
ATCCCTATATTTTCCGTTTGGAGAAAAATACTTTCTCTGCTAGACCTTT
TTTCACCAGCGATT

Sequence 3355

step.1001f07.cons.ok

20 ACTTGAAAAATCAAATGCAGATCGTGACATTTTCTATTTAAAAATCTATTGA
TAACAATATTTCGTGAATATCATATAGCAGAATAATCTAAACTCAAATCTTT
TGCTTTTTTAAATTTTTTAAAAATGAAAGAAACGAATTTACTATTTCATACTTG
CCCCCTAAAAACAACGTCATATAGCTGTTTTTAGGGGGTTTTTAAATCCCAA
GATAGTGCATTTTCATCTATTATAGACCTCTACTATTTATATACTGTTTACC
25 TACTATTTCATAGATAAAAAACGAATCCTATCCATAGATCAACATTGACAAAT
AACTAAAGCAATTCTATACAATAACGTATCTCAAGGTTTCGTTGTTTTTTTT
AGGTCTGTATAAATGTGTCATATCTATTTCTTCCAATTTTAAATAATTTTAT
TTTGTTGTCTAAAGTACCTCCGTATCCTGTTAAATTTCCATCTTTACCAAC
CACTCGATGACATGGTATAAGTATAGAAATAGGATTACTCCCAACAGCACC
30 TCCTACAGCCTGAGCAGACATTGCTGGCTTATTCATTGCTCGTCTACTGT
TTTAGCAATCTCTCCATATGTTTAAAGTTCTCTATAACGAATTGCTCGAAG
TTCATTCCATACCTTCTGTGAAAAATCAGTGCCTTTTGGTGCTAACGGTAC
TTTTATCTCAGGATAATTCCTTTTAAAAATACCTCTCTTAACCACGCTTTAGT
ATCTTTAAAAACATCTAACGTATCCTTTGTTTTAATTCCTTTCAAATGTTG
35 ATTAGGTATAAGACATGTGACAACGAGATGCCATCGTTTGTAAATTAATTG
CAATAAGCCAACGGGCGAGTCATACATTGATTGATACATATTTCTCCAACCT
CCTTAAATGTGTATATCATTTCTAGTCATATCCTTATAAACTAACGCATTCT
ATTCTGAACATAAATTTTATAATATAGATAATTGCTATATTTATGAAATTGG
ATAAACATTGCAATAGAATGTGATACACCATTAACGATAACTAAAAATTGAA
40 AAATGTCATCATTTTATATATTAAGTCAATTATCGACAGGTCATACTTTAAG
AAAGTCTGGAACCTTTGATTACGTCCCAGACTGGTCAAGATATAGAATTTA
AAGTCCAACTCACCTTTTGTATTAGATGTCATTATATTTATGAAAAATTAT
AAAGTTCATACCTTTTATTTTTTGCATATTTCTAATGCGAGAATTTCTAAAA
CATTAATGAACGCAGGTTCTGCTTTTACTTTTATATATTGAAAAGTGAGGC
45 CGTATTTATCTGCAAGTAAATGTCTTTTAAATGATTGGAATTAAGTCCTTAT
TTTGAGCATAAGCTTTTCTAAATTTAGTCGTACCTGTTTCTGAACGTGAAT
TGTATTGTCTTAATTTTGTATACAAATCACTATTTATAGGTTGATGTGTGG
ATGAATATCCTCCAGTTTGACCTACAAACACTGGTTTTTTCTCATCGAAAA
TCACATAGACACCAAACCTATCCTTAACTTTTCCCTATCCTCTGGTGTA
50 ACGCCAAAGCCTTTTTATGATCGTACTTAAATAATTTTTTTAGTTTCTTCT
GATTATCTTCTAGTAACCTATTAAATTTTTTTGCTACCCTTCATACATTTAT
TCCTCCCAATTAGCAATAAACTTATATACATACATTTTACATGTAGTATAT
CACTTTTTTAAACTTTTAAATCATTTTCAGTTTAAATAATTTTCACTATTAC
CTCAGGTACTAAAAAACCTAAGATTGAAAACACAAAACCATTTAAAACGTA
55 GCCTAAAATACTCCATGTGTGATTATAGCTCATCTGTAATTGTGTCTCGT
TTGAGCTATACGGTCACGTTCAAATCCATGCACGAGTCCCGCAACTACTGC
TGCAATAATCCCTGATGCATGGAATAATTCAGCAATCAAATAAGTGACAAA
CGGCGTCAGCAATTGTATAAATGTAAACATGTTAATGTTCTCAATGCCACG
TCTCATTAATGTTAAACGAAATCTCACTAAAGCCATGCCTATAATTAGACC

AACAAATAGCACCTCCAATAGAGGCAATCAAAAAATTGTTGAATCGCATCAAA
 AATGGAAAAAGTTCCAGTAATCAATACACCTACGGCAATTTTAAATGAAAT
 GATACCTGCAGCATCATTAAGTAAAGATTACCTTCAAGAATGGTCATTGA
 ACCTTTGGTAAAACCTTACCTTTCGTGATTGCTTGTACCGCAACAGCGTC
 5 TGTAGGACAAAGTATAGCCGCAATTGCGAAAAGCAGCACCTATAGGTAATT
 AGGCCAAATCCAATGAATAAACAATCCTACGCCCTATAACGGTAGTTATGAC
 AAGACCTAGAGCCATCATCATAACAGGTTTGATATACTTCTTAAATGTAC
 TCTTGAACATTAAACACCTCAACGAATAGTAGTGGTGCAATCAATGTAAC
 10 CATAAATAATTCAAGATCAAAATTGAATTCAACTGGAATGGGAGTGAGGTA
 TAAATCATCCCAAGGAAAAATTTGGATAAATGCAAGAGGTACTTTAGGTAT
 GAAAGTATGTACAAAGGAGCTCACAATAACTAGTGGGACAAATATGAGTAG
 AGTTTCAAATATCTCCACACTTTCACCTCTTTATTAATTTTAAATTGCAAA
 TTTAGTCATCAAAACAATAAAAAAGTTGCAAGGTATCTACTGCAACACTGA
 AGAAGTGAAATAATAACAATAGAGTCATTCTGTAAAGACATAACTATTAA
 15 CACTCAATCATGCCACAAATATACTATTCAAACTCAATTATTTCTTTAT
 CAGTGAAATCAAAACATGAACATTTAAATTAACCTATGAAGCTCGCTTCTTC
 GAACGAGAAAAATGCTTATGTGTAATTTCAATTGAAAAAATTAGTGACAACTG
 TCATTCAACAATTATTTCTGTTTACAATTTTATCATTTATTAACCAGATTCT
 TAAATAAATAGCTTTTATATCATTAATATCTTACCTCAGATATACAAAAA
 20 ACTACTAGCATATAAAAGATATATCTTTTATATGCTAGTAGTTTAAAGCT
 TAAACAATCAATCGTCCATTTTATTAATATATCTTTTCTTTATAGTATA
 ATTTTCTCGTCTTTCTTCTCGTAGTTTGCACGTTTTTTTAATTCTTCTG
 GCGTATTAAACTCTTTTGGCTTAATAAGATAGTAGATTGTCTAGATGCAT
 CGTCTTCTGTACATATATGCACCATTACGGTAGGCCGCTCTACTTATCA
 25 AATGCATTCCAACCGGAGATGTAAGATTGATAAAAACTAGTGATAATAATA
 ATCTGACACTGAAAAAACCTGAAACTATGCACTCATACATCAACAAATT
 TGAATCTTTTATAGATGATTTCAACCTTACTGTAAATCAAGAGATGAAATG
 GGCAAT

30

Sequence 3356

step.1001g01.cons.ok

TGGGCTGTTTCCCTTTCGAACACGGACCTTATCACCCATGTTCTGACTCCC
 AAGTTAAATTAATTGGCATTCCGAGTTTGTCTGAATTCGGTAACCCGAGAG
 35 GGGCCCTCGTCCAAACAGTGCTCTACCTCCAATAATCATCACTTGAGGCT
 AGCCCTAAAGCTATTTCCGGAGAGAACCAGCTATCTCCAAGTTCGATTGGAA
 TTCTCCGCTACCTCAGTTCATCCGCTCACTTTCAACGTAAGTCGGTTC
 GGCTCTCCATTAGTGTACCTGAACCTCAACCTGACCAAGGGTAGATCAC
 CTGGTTTCGGGTCTACGACCAAATACTCAACGCCCTATTACAGACTCGCTTT
 40 CGCTGCGGCTCCACATTTGCTGCTTAACCTTGCATCAGATCGTAACCTGCC
 GGTTCAATCTACAAAAGGCACGCCATCACCCATTAAACGGGCTCTGACTACT
 TGTAAGCACACGGTTTCAAGTTCTCTTTCACTCCCTTCCGGGGTACTTTT
 CACCTTTCCCTCACGGTACTGGTTCACTATCGGTCACTAGAGAGTATTTAG
 CCTTAGGAGATGGTCTCCAGATTCCGACGGAATTTACGTGCTCCGTCCG
 45 TACTCAGGATCCACTCAAGAGAGAATATGTTTTCGACTACAGGATTATTAC
 CTTCTTTGATTATCTTTCCAGATGATTGCTCTAACATGTTCTTTTGTAAC
 TCCGTATAGAGTGTCTTACAACCCCAACAAGCAAGCTTGTGGTTTGGGCT
 ATTCCCGTTTCGCTCGCCGCTACTCAGGGAATCGATTTTCTTTCTCTTCC
 TCCGGGTACTAAGATGTTTTCAGTTCTCCGGGTCTGCCTTCTGACATGCTAT
 50 GAATTCACATGTCGATAACATGACATAACTCATGCTGGGTTCCTCCCATTCG
 GAAATCTCTGGATCAACGCTTACTTACAGCTACCCAAAGCATATCGTCGTT
 AGTAACGTCTCTTCACTCGGCTTCTAGTGCCAAGGCATCCACCGTGCGCCCTT
 AATAACTTAATCTATGTTTCCACCATATTTTGAATTGTTATTCAAAATAAA
 TAGCTAAAACCTAGTTATTAATCTTGTGAGTGTCTTTTGAACACTAGCGAT
 55 TATTTATGAATTCAGCTTATTTAAACTCTATTCACTCGGTTTTGCTTGG
 TAAAACTCTACTTACTTATCTAGTTTTCAATGTACAAAGAATGTTAATAAA
 CATTCAAAACCTGAATACAATATGTCACGTTATCCCTCATCTTCGTAGAAGA
 TGTTCGGAATATATCTTAGAAAGGAGGTGATCCAGCCGCACCTTCCGATA
 CGGCTACCTTGTTACGACTTCACCCCAATCATTTGTCCACCTTCGACGGC

TAGCTCCAAATGGTTACTCCACCGGCTTCGGGTGTTACAAACTCTCGTGGT
 GTGACGGGCGGTGTGTACAAGACCCGGGAACGTATTCACCGTAGCATGCTG
 ATCTACGATTACTAGCGATTCCAGCTTCATATAGTCGAGTTGCAGACTACA
 ATCCGAAC TGAGAACAAC TTTATGGGATTTGCTTGACCTCGCGGTTTCGCT
 5 GCCCTTTGTATTGTCCATTGTAGCACGTGTGTAGCCCAAATCATAAGGGGC
 ATGATGATTTGACGTATCCCCACCTTCCTCCGGTTTGTACCGGCAGTCA
 ACTTAGAGTGCCCAACTTAATGATGGCAACTAAGCTTAAGGGTTGCGCTCG
 TTGCGGGACTTAACCCAACATCTCACGACACGAGCTGACGACAACCATGCA
 CCACCTGTTACTCTGTCCCCCGAAGGGGAAAACCTCTATCTCTAGAGGGGTC
 10 AGAGGATGTCAAGATTGGTAAGGTTCTTCGCGTTGCTTCGAATTAAACCA
 CATGCTCCACCGCTTTGTGCGGGTCCCCGTC AATTCCTTTGAGTTTCAACCT
 TGCGGTCTGTAATCCCCAGGCGAGTGCTTAATGCGTTAGCTGCAGCACTAA
 GGGGCGGAAACCCCTAACACTTAGCACTCATCGTTTACGGCGTGGACTAC
 CAGGGTATCTAATCCTGTTTGATCCCCACGCTTTCGCACATCAGCGTCAGT
 15 TACAGACAGAAAGTCGCCTTTCGCCACTGGTGTTCCTCCATATCTCTGCGC
 ATTTACCGCTACACATGGAATTCACCTTTCCTCTCTGCACTCAAGTTT
 CCAGTTTCCAATGACCTCCACGGTTGAGCCGTGGGCTTTCACATCAGACT
 TAAAAAACCGCCTACGCGCGCTTTACGCCCAATAATTCGGGATAACGCTTG
 CCACCTACGTATTACCGCGGCTGCTGGCACGTAGTTAGCCGTGGCTTCTG
 20 ATTAGGTACCGTCAAGACGTGCATAGTTACTTACACATTTGTTCTTCCCTA
 ATAACAGAGTTTACGATCCGAAGACCTTCATCACTCAGCGGCGTTGCTC
 CGTCAGGCTTTCGCCCATTGCGGAAGATTCCCTACTGCTGCCTCCCGTAGG
 AGTCTGGACCGTGTCTCAGTTCCAGTGTGGCCGATCACCCTCTCAGGTCCG
 CTACGCATCGTCGCCTTGGTAAGCCGTTACCTTACCAACTAGCTAATGCGG
 25 CCGGATCCATCTATAAGTGACAGCAAAACCGCTTTCACTATTGAACCAT
 GCGGTTCAATATATTATCCGGTATTAGCTCCGGTTTCCCGAAGTTATCCCA
 GTCTTATAGGTAGGTTATCCACGTGTTACTCACCCGTCCGCCGCTAACGTC
 AGAGGAGCAAGCTCCTCGTCTGTTCGCTCGACTTGCATGTATTAGGCACGC
 CGCCAGCGTTTCATCCTGAGCCAGGATCAAACTCTCCATAAAAAATTTATGAT
 30 GTTTGATTAGCTCATAAATACTAAATAGTTTGTAAACGTTTCGTTACTGTTT
 ATTGGAATTAACGTTGACATATTGTCTATTCAAGTTTCAATGTTCAATTAGT
 AAAATCAATCTTTTTTATTGTACTATATTGTTTTAAGAAAAGTCAATAACT
 AATTGATATTACTTAATTAATTTTAAAAGCTTTTAAAGATAAGTCTTTTTC
 TATATTCCTTATAATCATCCATTAAAGTTAATTCATAACCTACCTCATCAA
 35 ATACTTCAGGTTCAATCTTTTGTATTTTATATAATCCCCCAAAAATAAAT
 ATCTCTG

Sequence 3357

40 step.1001h02.cons.ok
 TATAACTTCCGTGAAACTTTAGCAGTTTTGCATCATCTTGTGTAATGGATC
 CAGTGATTGGATTAGCTAAACTTTGTTCAATAGTTCCACGCAAAAAGTCAC
 TATTTGCTTTTAAAAATTCATTTTCATCAAGATTTTATCTAATCTTCCG
 45 AAATATGATTATTTGTATTAAACCATGTATCCCCGTCCCTTCTATTAAATAC
 ACGTCTCTTTGATATCTTTTATCTCTTTTCAATTTGTTTAAAGTATTCTTCT
 GCATCTGTTTCAGATAGGTTTTCGCTCTTTGATTAAACACATTTTAAATCGCT
 TGATGAACATCCTTTGCCATTTTACTTTTCATCACCACATACATAAATAGTA
 GCGCCATTTTCAATCCATCGATTAAATTGTTCACTATTTTCTGCAATTTA
 50 TGTTGCACATACACTTTTTTATCAGTATCTCTAGAAAAAGCAACATCTAAT
 TTTGATAAAGTTCCATCTTCAAGCCATTCTTGCCATTCCGTTTGATACAGA
 AAATCTGTAGTGAAGTGTGATCTCCAAAGAATAACCATGTATTTCTTCA
 AAACCTAGTTTCTCTCGTCTTTCGATATAGGATCTAAACGGTGCAACACCT
 GTCCCAGGACCTATCATAATCACAGGTGTTGATTTCATCTTGCGGAAACTTA
 55 AAATTCGGATTTCTGTTTTAAATAGATAGGAATTGTATCGCCCTCTTGTATT
 CTCTCTGCAAATTGTACTGAACAAACACCTGAACGTTCCCGACCGTGTGCT
 TGATATCTAAGTCTCAACAGTAATGTGAACCTCATCTGGTGTGCTTTA
 TAACTACTAGATATTGAGTACTCTCTAGGTGGTAACCTTTCTTAATAATTGA
 TGTAATTTTCAGGTGTAGTTCTGTGCTGCGAAGTCATTTAATAAGTCA
 ATCAAATCCCTTCCCTCAACGTAGTTTGAATCCATCTTTATCTTGAATT

TTTTCAGAAAGCTCTTCATTATCAAAAAATATCGCAGCATTTCCTATTAAAC
GGTTTTGTAAATTTAGTAATTTCAAAATGCGATGTTAATGCCTCTTCAAGA
TTTAAAGTATCTCCATCCTCGTTAATTAAGTCTACGATAGCTGGATCATTTTGAGGCAAG
5 CCAATGTGCTAATAAGTAAGTCTACGATAGCTGGATCATTTTGAGGCAAG
ACAACACACAATCTCCCGGTTTCATATTCTCACCACAAATATCAAGTAAT
AACTCAATATGTCTCGTCTCTTTGTCTGAACCTCTACCATTCAAATTAATA
TTGGTCAACTTTCAGCATCGTAAGGATTAGACTTACTATATTTCCTTTTCT
TTTGCTGATTTTATAGATTTCGCTAACAACTTGTTCACTTTCAGTACCAGCC
GGTGTGAATCAATCGTATTAATAACATTGGCCATCCACTTTTCAGCGTCT
10 TCTTCATAATCAACATCGCAATCAGTACGATGATAAAGTCTTTCTGCCCCCT
AATTTCAGCCAAACGGTTATCAAAATCTTTACCAGTTTGACAGAAGAATTCA
TAGGTTTGATCTCCTAATGCTAATACTGAAAATCTCACCCATCCAATTTT
GGCGCTTTACGTCCGTGGATGTATTCATGTAATTCAACAGCGTTATCTGGT
GGATCTCCTTCGCCATGTGTAGATGTGATAATAAATAAATCTTCTACTTTC
15 TTCAAATTCCTTAGGTTTAAAGTCATCCATTGATTTGAGCGTAACGTCATTT
CCAATATCAGATAAACGTTGTTCAAATATTTCTGCAAGTCCTTGCGCATTA
CCAGACTCAGATCCATATAAAATTTGTTATCGATCTAGCTTCGGGCTCAACT
GAAGGTTCCCTTTTCATGTAGCATCGCTTCTGTATTGTTCATCTGAGTTATGT
TGTTCTACAGAATCTGTAGATGTATTTGATTGTTGATTGCCATCAGATAG
20 CCACTTAACCATACTTGTGGTTAGATGATAAAGTGTGAAGCAATTCATTG
ATTTGTTTAGCTTGCCCTTCCGTAAAAGGACTATTGGTAACACTTAAGTTC
AATCTTATCCACCTCTTATACCTTTGTTATTCCGATATTTGTCAATTATTTAT
ATATGTCTATTTTGAAAGTTATTAATTATATCTTCAATTCAGTATGTAC
CATTTGTTATTTTCAGTAAATGTAATACTCAATTGATAACTAAAATCGAAA
25 AGACAATTTCATAATATCAAGGGCAATACAACTTTTATTACTCAAAATTA
TCGTTACTAAATAAAGCTGTATCAGCTTTTAAACAACAACCTCTTTATTTCT
TATCGGAATACTATGTTTTGAAATTAATAATTTTGTACTATCTCTTATGGT
TTATCAGCTACATGTAATCCACATTCAGTCTTACTAGAATTGGACCAACGA
CCAGCACGTGAATCATTAGAATCAAAATACGGGTGATGTACATGGAATACAA
30 CCAATACTTGGATAAATTTGATCATGTAATTCATTATATGGTAAATCCTTA
CCAGTATATAAGACCATACTTCTTCTTCTGTCCAATAGATTAAGGGACAC
ACTTTAATTTGACTTAAATCTTTCATCTTTGTTAATGAAATTTGTATGTGCT
CGTGTGTTGGTATGTGCTCGTCTAAGACCTGATATCCAAGCTACAGCACCA
GATAATACGTCTTCTAGTGGTTAATCTTGCGTATGTAGCAACATTGGTTA
35 GGATCATCTCTCCATAAAGCAGGATTATATTTCTCACCTTGTTCCTCAAC
GTAAGTTCGGGTTTTTTTCATTTTAATGCGTAATTGCGGATATTTATCTTTC
ACCTTATCTATTAAGTCATATGTTCTTGAATAATGTAATCAGTGTCTAAA
AATACAATTTGTGCGTCCGGTTAATTTGAGAAATCAAGTCAATCAAAACC
ATACTCTCAGCACCAAACTGCAAGAATAAACAATATCATTTCATAAGTT
40 TGATATGCCCATTTTAAATCTCATAAGCACCTTTTGTCTCATCATTAATA
TCTAATTCATTTATAAATGGATCGTTTTGAAATTTGTCATAAGTAATCTT
TCTATAGACATTAGCGGAGCCCCCTCTCACCTTATCAATCCATCCAAATCA
ACAAAACATATCAGAATTGTCAGAATAATTATGTTTAATACTTTAACTACT
TTTATAACGAAAGTCAAGGTAACTCGAGAAATTTATCGAAATAAAGATAAGA
45 TTAATTTAAATTTTGTGCTATATGCTCATAAAAAATTTAAAGAGCCTAC
ACATATGATATGGCAGACTCTTTAATAATTTCTTTTACATTCGTTTATAGT
GACTTCATTCAAAATATTACAAATGGAACAAGATTATTTTATTCTTTAATC
CAAGAAGAAATATAGGCATTTATTTAATAATTGCTTTATTAATTTTCCAT
TTGGTTAGGATTATCACAAGGTAAAAATCGTTTAAACAATATTTCCCTGTTG
50 ATCTACTACAAATTTAGTAAATTTCCATTTAATTTGCGAACCATAATCC
TGTTGTTTACATTTTAAATGTTGTACAAAGGTTTCATGCTCTCCATT
AACATTAATCTTAGCATGGATGGGGAAGTAATACCAAACTTATATTTATA
TACTCGATAAATATCTTTGATTAAATTTCTGGCTCCTGATTATTAATAATCAT
GCAAGGGAAGCTCAAAATTTCAAGACCATACTTATGATATTTTTTATAAAG
55 CATTTCTAATTTATTAATTTGATCATTTCAGAGTACAATTCGTTGCAGTATT
TACGATAATGAGCACTTTCCCTTTATATCTTTTATAGTAAATAGGTAGAGCC
ATCATAATCTTTACCGCAATATCATAAATACTCATTTAGATACACCTCTA
CATCATTTGCCAATTTCAATAAATCTATAACGCCACTAGAAGCTCCAGGAAA
TGCAAAAATATTTTTATTTAAGTCTTGTGCAGTAAGTTTTTGATTAAATGAT

AAAGACTAGGAGATTAACCAAATCACCAGCGTCATTACCATAAATCTCTGC
TCCTACTAAACGTTTGTGACTATCTA

5 Sequence 3358
 step.1001h07.cons.ok
TAGGCTAGTGGTACTAAAACACGGTCTTTATAATGTTTGAATCTGTTGTAA
TAAAACTATCTTCTCTATCTGCAAATCTTTAGTTTCAGAGGTATCCTCC
ATAAATGACCTAAATATAGGTAACCTCTCTTCAGATAAAAAGCGGACTTTA
10 ACTCCATTTTCTTAACTTTTTTAGTATTACGCTTTCTTAAACCATCCATG
TTTTTTAAAACATCATTAGCACTTTGTGTTGCTAAATTTAGAACAGAATGA
TATCGGATTTGTAATACAGGATCAAATCCTTTGTGGAATCCTTCGTGTTTA
TATCCTAAACTCTCTAATTCATCAAAAATCCAATCATGACCTGCATTTCCA
GTTATTTCTCCCTCATGATTTAAATATTGATATGGAAGGTATGGGTCAACT
15 CTTAAATATAACAATTATATTTTTTACATATTTACTCAATTCATTAAAG
AAAAAATGTACAAGCTCTTTATTATTATAATCTATTACTGGACCGCGATTG
GAATAAAAATATTTAAATATTTTCATTACAGGAACAGCTGTTAATAAACAA
GCTGCAATCATTCTCGTTATCATTATTTTAAATTTCCAACATAAATGTGACTCG
GTACCTTCAGCAACCTTTAATTCGTAATTACCTTCCATTTGTGTAATAATGA
20 CTATATGTCATACGATCAGTAAAGTCACTAAATTCCTTAGCTGTCAAATTC
GTAAACTTCATCTTCATAACTCCCATTTATATCTTTATATCTATTTCTTTT
TATTTTTTAGAAAAATAGGTGCTTTTACTTCACCTTTTATATTTAACATTTG
GTTTCATTTAATACTATTTATATTTCTAATCTAAACAAGTATACAATAGTTG
TAGTGAAACTAAAAATATTATCCTGCAAACTTTTATATAAAATTTAAAGT
25 TAAATCTCGTTTATTCTGTTTATGTATAAAATAATATGGGCTAAACACTAA
TGTCGCCCATATCACTAATTATTTACTATTATTTTAAATTCATTTGCAAGTG
CTTTACCAATGTCTTCACTTGATTGGGGATTTTGTCCAGTGATAAATTGAC
CGTCTTTTTCTACGTGAGATGTAAAGTCATCTTTCACTACAAAATTTGCAC
CTTGCTCTTCTAATTTAGATTGAGTTAAAAATGGTACTTTATTTTCAAATC
30 CCATTGCTTTTCTTCACTATCAGTAAATGAAGTTATTTTGACACCATCTA
CTAGATAGTGATTTATTTGCATCTTTTACACCTACAAACGCCTAGGTCCAT
GACATCTGAAGAGATTATTTTATTTACTATTTTAAATTTGAAGTAAAAATAT
CAGCTAATTTCTCATTATTGGCAAAATCGTATACAGTACCATGTCCACCTG
GTAGATAAATAGCGTCATACTCATCTGCATTTACATTTTCGATACTAGGTG
35 TATCGTTTAAAGTGTGACACGAATTTAGCATACTGATTTCAGTGATTTCATTAG
AAACAGAAATTAGGATCAAGATTTACTTTTCCACCTTTAATAGAAATAACAT
CAACATTGATACCTTCTTCAGTCAATATATTATATGGTGCTCCAGCTTCTT
CTAACCATAAATCCAGTTTCTGTACCGTCTGTAAATTGACTTGTACTTGTFA
AAACAAATAAAACTTTTCTTACTCATGAATATATACACCTCTTCAAATTAAT
40 TGTGATTATTTATTTACCCACATTTAGTAGTAGATTATTCATAGTGATCAT
TAATTTTTTGAAGTTACTATTAGTATAGTAGTTAATTCATAGAGTAATCAT
TGTTTTTGATATGTATATTTTTCATCTATTTTCTTAACTTTGATTTAACTT
AATATTAATTTTATAAAGAATTCAACGTAGTTCTTATTGATTTGATATAA
GTGATAAATCTTTTCTTGAATCTATTTCAATACGTTTACAATTTCACTA
45 CCAATTACAATACCATCTGCAACGACGCTATATCTTTAACATGTTTCAGGA
TTTTTTGATACCAATCCAGCAACCACAGGAATTTTGAACCTTTTTTTTATA
TATTCAATTTTCTCTTTAAATCTGGATGGAACCTCCCACTGTTACCTGTT
GTGGCATTCATGTTTACCGTGTAATAAATCCTTCTGAGTTCTTTGCAATT
TGCATAATCCTAGCATCACTTGCGGTGATGGCAATTAACGATATTATTTTA
50 ACAGAATGATGATAAAAAATCTTTTTTAAACTTTTTTGTAAAGTTTCGTAAGGT
AAATCTGGAATAATTAACCATAAACACCAGCTTCATCACACTTATCCAAA
AATAATCTTCTCCATAAGCACCTAGAATATTATAAATAGTCATTAATACA
TACTTAGATGAAATAGTATTTTATTTTATTAATTCATCAAAAATGAAT
TTAATGTTTGAACCTCGTCAATAGCGTTGCGCCCTGCTTTTCATGATTATA
55 GGTCCATCTGCAACAGGATCAGAAAATGGCACACCAATTTCAACAATGTCT
GCTCCATTCTCAGTTAATGTTTTTAAATGATGAATAAAATTTAAATCACCC
ATAATATACGGAATGAATAATTTACTCATTTTGTTCACCACCGTTTTCTTT
GTATTGTTTAAATGTTTTCATATCTTTATCTCCACGACCTGAAATAGTCAC
AACAAATATTTCTTTTTTCATCCATATTTGGCGCTAATTTTCAACATAACT

CAATGCATGTGCACTTTCAATTGCTGGAATGATACCTTCAACTTTTGAGAA
TGTTATAAGTGCTTCCATAGCTTCATTATCTGTAGCACTTACATATGATAC
ACGACCAATATCATTATAATACGAATGTTTCAGGTCCAATACCTGGATAATC
TAGTCCCGCTGATATAGAGTGTGCCAATTC AATTTGTCCATCATCATTTTG
5 AATAAGGTACATTTTGAACCATGTAATACACCTGGTTTACCTTTCCCTAT
AGCTAAAGCATGATTATGCGTATGACTTCCTTTTCCCGCAGCTTCTACCCC
ATATAATTTAACATCATCTTGTATAAATGGATAGAACGTACCTATTGAATT
GGATCCTCCACCAACACACGCGACTAACGCATCTGGAAGTCGTCCCTTCTTT
ACTT

10

Sequence 3359

step.1001h09.cons.ok

CACAACCTGAAGCTTGTAAACCTGGCTTAATAATTGTATTTGCTCCTACTTG
15 TTGATCGTATTGTTTCATATAAATGATGTTTAGATGCTATTGTTGGGTGTTG
TAATAACTTGATAAATGTATCATGTACATCTATATGCTGTAATCGTTTTT
TGATGTATTATATTCCTTTTCTCCTTCTAAAAATATACACTGGTGCCTC
ATCAGCAAGAGGTTGCACAGGAATGTCTGCATACACTTCATCATCGTATGT
AAGCACGAATCTATCAGTGTCTGTTACTTCACCGATTACTGCACTATCTAG
20 TTCGTGTTTATTGAATAAATCTAAAACTTTTGTCTGTACCTTTTCTAC
AACTCTTTAAGCCCATGTTTTTATAGGTTCTGTGCAAGTAAAAAATA
TAGCTAACCCTAATTTATCATGTCTAGTGTTCGCTTAACTTGCTAGCATGA
TGCTAATTTTCGTGGCATGGCGAAAAATCCGTAGATCTGAAGAGACCTGCGGT
TCTTTTTATATAGAGCGTAAATACATTTCAATACCTTTTAAAGTATTCTTTG
25 CTGTATTGATATTTTGATACCTTGTCTTTCTTATTTAATATGACAGTGAT
CTTGCTCAATGAGGTTATTCAGATATTTTCGATGTACAATAACAGTCAGGTT
TAAGTTTAAAGCTTTAATTACTTTAGCCATTGTACCTTCGTTGAAGGTG
CCTGATCTGTAATTACCTTTTGAGGTTTACCAAATTGTTTAAATGAGACGTT
TGATAAACGCATATGCTGAATGATTATCTCGTTGCTTACGCAACCAAATAT
30 CTAATGTATATCCCTCTGCATCAATGGCACGATATAAATAGCTCCATTTTC
CTTTTATTTTGATGTACGTCATCAATACGCCATTTGTAATTAGCTTTTT
ATGCTTTTTCTTCCAAATTTGATATAAAATTTGGGGCATATTCTTGAACCCA
ACGGTAGACCGTTGAATGATGAACGTTTACACCACGTTCCCTTAATATTTTC
AGATATATACAGATAACTCAATGCATATCTTAGATAGTAGCCAACGGCTAC
35 AGTGATAACATCCTTGTAAATTTGTTTATATCTGAAATAGTTCATACAGAA
GACTCCTTTTGTAAATTTACTATAGATTCAACTTTGCAACAGAACCT
ATATTTTGTATATAATTATTAATTAATTGCCTTTGTGTTAGTATTCCTTGG
ATTACTAACACTTTGGTTAGCTGTTCTAAGTGATACAGGTGCAGCAATATT
AGTTATATTAAATTCATTACGTTTATTAAAAAATAAAAAAGCCTAAATAC
40 TCTCTCGATTGTGAGAAAGTATTAGGCTTTTGAATTAATAAATTTCTT
ACTTCTTATGTTTATTTAACAATTAGTACTGGGACAGAATCTTTTATAGCA
ATCTTATGTGTAACATCTCCTAAACATGTTTAAATTCAAATCGACTCTA
TGATTACTCATAATTACTAAATCAAATTCACCATTATTTATATCTCCAGAA
GTAATTTTCTTAAGATTGTATCTTTAATATGCCCAAAATCTACATCTATT
45 TTATATTGAATATCTCTTTTCTAATTCACGTAAAAATGGAGTGAGTTT
TTCTCTTTTCTCTAATAATATCTTCTTTATGCTTATTATAATACCTTGACA
CTAATTGCAAGATCATTTTCTGTACAAACGTGATAAATTACAACAACCTGAA
TCTTCATCGGTGACTTTTTCTAAATAGTCAGGGATAGCACTAAATCATTT
TCGAAATCATAGGGTAATAATATATTTTGAACATCTACACCAGCCTCCTA
50 TTACTTATTATTTAGATGTCATTACTTATATATACCCTAAATATGTATGC
TTATTCTCATATTCAAATTATTCAAATATATTCCTTATCATGTATTAGATT
CTATTTTTTATATTATTTTGTAAATCTTTGTATAAACCTATTTATTTATA
AAGCGCTCTAATACATACCTAAATATCTTCAACTTATAAACTACTTTTAT
TTCAATGCCCTTTTCCAAACTTCTTTTATGAGATATTGTGGTTATTACCTA
55 ACTCTCATTTAATGTTATTATTTTATTTATACATAGTAAATGTAACCTGT
AAGACATTTCTACCTAACATGTTTTGTATTACCTGACATTTACAATTTAC
ACATTTTCGCTATTATTACGCAATTAATTTGATTTTATAAATTATGATGACT
ATAATCTATTTCTAATTAATCTCTCTTTTAGGCATAAAAAATCTACTTAT
AATTAAGCATCCACTATTTCAAAATTAATTTGTAAATTGAGATTTATAAGT

GATTTTTATACAGTTATTTAAAAATCTATTTTCATAATCAAAAATAATAAAGG
 GGGTAATAAAAAATGAAAAAGAAAAACGTTTAAATCTTATCTTAACTGTTA
 TACAACAGAATCAGTTTAAATAAAAAACAACAAATAGTGGATTACATGGCAA
 GACATTTTGGAGTTTACTATAGCTTGACAACATTTCTCGTGACTTACAAG
 5 AATTAGAAATTTACAAAATCCCTGTTGAAAATAAAAAAGTATATTTACAAGA
 AAATAAATCAACAAATCAATTAAGTGCAAAAAACAATTAGAAAATTTTA
 GTGATGAGATTATTGAATTTATAACGCTAAATAACTATGTCTTAATAAAAA
 CATCTCCTGGCTTTGCTCAAAGTATAAGTTATTACATAGATCAATTACAAA
 TGAAAGAAATATTAGGAATTTATTGGAGGTAACGATACTTTGATGATTTTGA
 10 CTTCTTCAAATGAAATAGCAGAAATTTGTTTGTATCAATTATTCCCTTAAA
 AATTAAATGATTATTAATAGGAATGATTACTAATACAAAAAATGAGTA
 AAAAATCTTTTGTCTGAATAAGATAAGGTTTAAACAATGAGATTTT

15 Sequence 3360
 step.1002a01.cons.ok
 AACACAACCTTATTAACTTTTGATAAGATTTTATAGTATTAAAGTTCCAT
 TCTTTACGAGTATTCATAGCTTCTAAATAATACCCAGTTATAATATCAAAT
 GCTTTCTCTCTAAGCTCTCTCTCTATTTGAAACATGATAAGGAGCACCA
 20 CCTTTAGTAACCCACTTATCTTTTATAATTAGATAAATAGGCAGTATAT
 AAAGAGATAATGGTCTCTTTGAAACCGAAAAATTTATACTCTCTTTTAAA
 GAATCTTCACTAGCCACTATAGAAAATAGATGAGCGTTAAAACTGAGGGG
 TTATAACTTTCTGATTTAGGTGGAGCACCACCACCAGAAATTATCAATA
 GGCAAACCATCTTTATATCTTACTCCAAATAATTTCTCCGATTTTAGCGCTC
 ATCTCATCAGCATCAAATCTTATACTTTTATCATTTTCTTTTCAAACGT
 25 CTTAAATACCTCTCCTGATCTTTTCAATACCTTTACGTGTTACTTTTATC
 GATCCACTAGGGAATTTGAATTCACCCACTGTAAATCAGTCAAAATTTT
 TTTATTTGTCTTTAAATAAAAAAATGATGATAAGAAATAAACTGGAAAA
 GAAAAATCTGATACAAAAATTTAACGAACTAAAAATTGATTTTAATATC
 30 TCATAAATACACTCCAAAACCTACCTCCAAAATTTTAATTGACTTTTAGAT
 TATTACACTAAGAAATCATTAATAGTAAAAACATATATTTATTCTAATTAC
 TTTATTGACGTTGATCCTCTGAACCCTTAACAAATCCAAAACCTGTGCAAT
 GGTGGTTTAATAAECTTACGCTATGCCGACATTCGTTTCAATCTGTTTGAT
 AAGTTAAATTTTCATATATAGTCGTGAATTTCAATGTTCTCTTTAAACAA
 35 GGATCGTGAAAAATATGCCATAAAAATCAAACTAAATGATTCATTAGTCTTT
 AATTATTTTCTGTTACTTATAGGTGTGATAACCTTTGTGCTTATTTATTT
 TTAGGTGAAGTTAATGTTAATATAGATTTATAACGTTTAAATGTCATGTTT
 TAAATTGTATTTTCATGTTTGCATCTCTCTTTTTFATTTATCTCTTAGTC
 AAGATTTAAATCTTAACTAAGAGATTTTGTATTCTTTACATAATAAAGT
 40 CTTATGGGTAGTTGAATTGTATACTCTATTCAATTTACTGAATTATATCAAT
 AATTTTATTCATCTTATCTTTATCACGTGATACTACGCCTACAACCTTTT
 ATAACCTTTGTAATTTTTTTTGATATAAGTTTCATCTTTCAAGTCTTGTAGT
 CAAAGATTGAGGAGTCAAAGTAATACCTTTATTTGCTTTGACATTAGCTAA
 TATAGTTTTCATAATAATCAACATAATCAATATGTAACCTTATCTTTCAAGGA
 45 ATAATTTTAAAGAAGATCATATGTATCACATGGCGGTGTTGTATATAAAT
 ATTCTCATTTTTTAAAGATTTTATACTTACTTGGTTAGTGTTTTTAACTT
 ATTTTCATCACCATAAGCCACAATAAAGTCCTCAGTGTAATTTCTTTGTGT
 ATAGAGCGAGTTATTTTTTAAATGTGTAATATCACCAATAACAACATCTAT
 ATTATCTTTATAAATTTCCCTCAAGTAAATGAAGTACTATTTTCAATAAC
 50 TAAACGACACTTTTCATTAATATAATCTTTTTCATGTAGCTTATATAAAGA
 AAAGCTAGGTATGACTCCTAATTTAATCTTTTATCTTTTAGACTCACTTAT
 TTTACTTTTAAATCTTCAAATTGCTCGTTTATTTTAAAAACGAATGATA
 AACAGCATTTCTTTTGTAGTTAAAAAATACCCCTAGAAGTTCTATAAAA
 TAACTGACAAATCAAAATAATCTTCAATACTTCTAATTTCTTTACTTAAGGC
 55 TGGTGTACTTATATTCAATTTCTCTAGAAGCTTTATTTAAGCTATTGTACTC
 GACAACCTTAAATGAAGTCCTTCATATAAATAAAAATCATAAAAATCCCCAT
 GGTATTAACATTTGTTTATACCCATATTAATACATTACTTATTCATATAT
 TTCAAGTAAAGGTTATCCTTCTATTGTTCCCAAAAATAGGCAACTAATCT
 CTTAATGAATCAAAATGACTGATAGGATGAGATAAATGAAAAAATAT

GGATTTTAACTATAGGTATGTTTGCCTTAGGTATGGATGCTTATATAGTAG
 CAGGATTAATACCTTCAATAAGTAAAAGTTTTAATAAAAGTAGCTCTGCTA
 TTGGGCAAGGAGTAACAGTTTTTACATTGTTTTCTCTATCTCTGCCCCA
 TTTTTTCAACAATATTAGCTAAATCCCCAGTTAAAAAATACTAATAATAG
 5 CATTACAGTATATTTACTTTAGCCAAATATTATAACCGCAATATCTATGAACT
 ACATGCTATATATCGTATCAAGAGCAATCGCTGGTTTAGGAGCTGGCGTAT
 TCTACCAATTGCAATAAGTGCAAGCAATCATTTAGTCTCCGAAAAGCATA
 AAGGAAAAGCAATCGCTTTTACAGTAGGCGGAATGAGTGTAGGAACGTTA
 TAGGAGTTCCCTCTCGGACTAGAAATTGCCAACATTTCTAATTGGCGATTG
 10 CAATGTTGGTTATTATTGTCAATTAGTTTTATTGCATTAATAAGCATATCTA
 TATTGATGCCATAAATTTAAAATAGAAAGCTCCTCCAAATTTAAAAGATCGTT
 TTCAATTATTTTTTAAACAAGCATGTACTAAGAGTTATTTCGGTTACATTAT
 GCGCTGCCATTGCTAGTTTAGGTTTGTATACTTATTTAGCCGATATTATTA
 AAACAAATACAGATACAAAAATTTAACTCATTACCTTACAGCGTGGGGAA
 15 TAGGCGGATTAAATAGGAAGTTTTTGGTATAGGATTTATATAGATAGATTTA
 AAAATACAAGATTTGTTATGCTAATTATTTTAATTTTACTAGCATTAAGTT
 TTGGTTTAATTCCTATTTCTATTAACTTGCTTATATTAGGTTTAATTCCTT
 TTATTTTATGGGAGCTATGGGATGGGCTACACAAGCTCCTCAACAACATA
 TATTATTGAAAAAACATCCTGAATATGGAGGCTCTGCTGTGCTTTAAATA
 20 GTT

Sequence 3361

step.1002a02.cons.ok

25 TTAATTAATTTTATTTATTTAATTAATATTTTGTAAAAATAAATACAATAA
 TACATATTTATGATATGTTTTTAAATGTAATAAAGTCAAAATAAAGGAGGA
 AATTTATGAAAAAGTTAATGATGAGCTTATTATGTTGCACAATATGTGGAA
 CAGTTTTAACTACAGGAAGTGTAATGCACAAGGTGATAGCTCAACTAACT
 CTGAGTCGTTAAAGGAACTTCAAAATGAAGGTATTGTGTCTAAATCAATTA
 30 CTGAACAACAATGGCAACAAATGAAAGCTCAAGAGCGTAAAGATGAAGCAG
 AATTTGAGAAAACAGCTGAAGTACAATGGCAGAAAACAAAAACAAGATC
 GAATAGATCGAGAAAATCGTTCAAGAAAAAGAAATTTCAATTTGAAAAAAG
 GTGATTTTTTTCATCACAAATAACGTAAGTTCAAAAGGGTTTACAGGTCATG
 CTGCAATATATACTGGAAGGAAAAGTTAAAGAAGCGCTTGATATGGAC
 35 AACCTGTGAGGGTAAAAAGTTTTAGTGATTGGAAAAAGAGTACTTTGAAAA
 AAAGAAAAGGAGTCCCAACATCGTTATATCAAGGTTTATCGAGCACC
 AAAATATAGAGGTAAAGCTGGAACATATGCGAAATCTCATTTTAACGGTG
 TACCTTACAGTATAACGACGAATCCATATTCTAAAAGCGTTACATACTGTT
 CGAAACTTGTGTTGGCAATCGTATTATTATGGTGCCGGACGATTATCAGTTC
 40 TTCCGGTTGTTACATCTCAATTTATTATTGAGCCATATAGTCTCAATAAAT
 ATATTCCATCAAAAGCAGTTAGGCTTATAAAAGAAGCTAAAAAATTAAGG
 TGTTAAATATATAATTTATACATTTTAATTACCCATAATCGTTAATTATGT
 TACTTTTCTTTGATTTTCATCGTTTATTAGAACACTTTGATAATACATCAT
 TATGAACCTCAGTACATATTATGTGATAATCGACCATTAATAAATACGATT
 45 AAATGATTAAAAAGAACAAATAAAAAAGCAAGCGACTAATAATATGATTATTA
 GTCGCTTTCAAACTGTCAATAAGCGCTCTTTTCAGATACTTTATTTATCTA
 ACTCTACTGTTTTATCTTTGGTATTTCTGGAAGAAATGCATACTTAATCT
 TTATAAATAATGAAGCTGGCTTTCCATTGATATCAATACGATAATAACCTA
 TATGCCACCTTCCTTTGATTATCAGATTTTCCATCTATAGAAATTTTATTTT
 50 GCTTTTTTAAAGAAATGATACGTTTTATTATCTTTAGCAACCTCTTTTAAAT
 TCTTAGTACTTTCATGCTTAACATAATTATTTAATTGCTTTTCACTTGAAT
 GACAGCCCCATGAATCAAAATAGAAAAAGTTTACAAAAATATAACAAATG
 CTAATAAATTTATTATACCAATAACATAGAATACTTTTTCATTAAAGTTC
 ACCTAGTTTTCAAAAAAATTTTTATAATTTATGATAACTTAAAGTGATA
 55 TTTTTTCTACATAATATATTTAATTTCAAGGAGATGTAAAAAAGTTGAAA
 AATTTTCGCAAACTAATTTTGGTGGGCATTTTAGTATCGGGGTCAGGGATA
 GCGAGTGTAACAACAAATATAACTCACGCAAAAGAAAGTCAGGATTCAACT
 CCTCAAAATATTAATTTAGTGGGAACGTATGATACTTCTCAAGTTGATTCC
 AAAACGATGAACAATTTAAAGAAATAGAAAAAGAGATAATAATTTCCAC

5 ATAAC TAAACATG GAAATAAAGTCGTTGTAG AAGACAAATTACCTAATCCA
 GAGAATAAAACTTCAAGTTATTCAGCTGATGGTAGTGCTGAAAAATAATACA
 AAAGTAATTAATTTCTCTGATTTTGTGGTAATATGGATGGGAAAGATGAT
 GGAATAATATCGGATGGGATAACCTTTTATAGTGGTAATCATATAACGGA
 10 CAACACGATGGTCAAAAAGTAAAAAAGGGACTCATGTACATTGTAATAGA
 TTTAACGGAACAAAATCTGATCATAGATACTGGTCAAAAAACATCCTAGA
 GCTTATGTAGATTTTATAAAAGTGATTGCTGGTATCACGCCAAAGCTTAT
 AAATGTTCTTCCTTGGGAAAAATGACTAAATGCGATGGTTTGAATAGTATT
 TATAGAAAAGGTGTCAAAGATTGCTCATCATGGAAAGGTAAACCCAAACAT
 15 AAAAACTGGCCTAAAACAGCATGGTATAGAAATTA AAAAATGAATAAAATC
 TTA AAAATATTAATAAC TTCTATTATTGTTATCATTATTACCTTAACAGTT
 TGGACTTTTAGTGTGATTACTTATCAGAAACACAAGAGTGAGAAAATCATC
 AATCACGTTATAGAACGTAAGGGTTGGGATAAAAAAATAAAAAATGAAAAA
 ATGAGTTTAAATATTATAATGGGATATGCTGAAAAAGATATTGTTTAA
 20 GATCAACCATATAGTGAGTATGAGTATAACGTGACACCAGCACCATTGGACA
 GATGATAAAGAAATATAAGGTGTGGGGGAAACAGATTTACAAAAGAAAGAC
 TCCTATTATAAAATATCTTTTAGAATCAGAACCTTACAGAAAATAATATAAA
 TGATAAACATAATAAAAAATAGAGATCATTTCATGGACTTCTAGTAAGATT
 AAGGTAGAAATACAGATGGTGGATTAGAATGGTCACCAGATCATAAATCAC
 25 TTATTTATAAAATATGATGCAACATTAGGTAGACAAAATAAATACTAATGACG
 TGTTAACTTTACTTCAAGCAACAGCTAAAAACTCAAATTTACGTTCAAATA
 TCAATAGTAATGAAAAACAGTTAGCAGAACGAGGGTCTAATGGGTATTCTA
 AATCTATAATTAGAGATGATGGCGAGAAATCTTATTTACTTAACTCAAATC
 CTATTCAAGTATTAGACTTAGTAGAACCAGATAATGGTTACGGTGGACGTC
 30 AAGTCAGTCATTCTAACGTTATATATAATGAAAAAAATTCTTCTATCGTAA
 ATGGTCAAGTTCAGAAGCTAATGGGGCATCCGCTTTTAATATTGATAAAG
 TTGTTAAAGCTAATGCGGCAAAATAATGGTATTATGGGTGTTATCTATAAGG
 CACAATTATACTTAGCACCATACAGTCCAAAAGGTTACATTGAAAAATTAG
 GCCAAAATTTAAGCAATACCAATAACGTGATTAAATGTTATTTTGTGCCTT
 35 CTGATAAAGTAAATCCTAGTATAACTGTAGGTAATTACGACCATCATACGG
 TATATTCTGGTGAAACATTTAAAAATACTATCAATGTAAATGATAATTATG
 GATTAAATACAGTAGCTTCTACAAGTGATAGTGCAATTACTATGACCAGAA
 ACAACAACGAGTTAGTAGGTCAGGCTCCTAATGTTACTAATAGCACAAATA
 AAATTGTAAGTTAAAGCCACAGATAAAAGTGGAATGAAAGTATGTGTTT
 40 CTTTCACAGTAAATATAAAACCATTAACGAGAAATATAGAATAACAACCTT
 CATCAAGTAATCAAACACCAAGTGAGAATTAGTAATATTCAAAACAATGCTA
 ACCTTTCAATTGAAGATCAAAATAGAGTAAATCTTCACTCAGCATGACTA
 AAATTTTAGGTACAAGAAATTAATGTCAATGAGTCAAATAATGACGTTTCGTA
 GTCAGGTTGTAAGTAAAGTAAATAGAAGTGGGAACAATGCTACAGTTAATG
 45 TTACAAC TACATTTCTGATGGTACA ACTAATACAATAACCGTTCCAGTTA
 AACATGTGTTATTAGAAGTTGTACCTACTACTAGAACAACAGTAAGAGGAC
 AACAAATTTCCAACCGGCAAGGAACCTCCCAAAATGATTCTTTAGTTTAA
 GAACGGGAGGTCCAGTTGATGCGAGAATAGTTTGGGTAAATAATCAGGGAC
 CCGATATAAATAGTAATCAAATGGTAGAGATTAAACATTACACGCTGAAA
 TATTCTTTGA

Sequence 3362

step.1002a06.cons.ok

50 GGTGCGATATATACTTTATTTCAACCATATTTAAATACTATGATGCACTAT
 TGGTTAATCAAATAAGATTATTATGAGAAAAATAACAATTA AAAATTACGTA
 TACATAAGGAGCGTTATCATGTTAGACAAAAATCAATTGGAAAAAGTATAAC
 CAAGAGCATTTGTATGAATATGAAAAATTAATGAGTAGTAATGAAAAGAAT
 GCTTTAGATGAAAAAGTAGATCAGTTAAATCTTGCAGAAATTCAGATTTA
 55 TATCAAGATTATATGTTAATAGAAAAACTATTGATGATGTATCTTCTGTA
 TCTGAAGTCAAATATGAAGTGAATCACGACTCAATGAAGAAGAACGACAT
 ACATATGAACAAAAAGGTTATGAGGCAATACGAAATGGTGAATTTGCTGTA
 TTATTGATGGCTGGAGGACAAGGTACGCGTTTAGGATATAAAGGGCCTAAA
 GGTCTTTTGAAATAGAGGGTACGAGTTTATTTGAACTTCAGGCGCGTCAA

CTGATTTCGTTTAAAAGAAGAAACCGGCCACACAATTAATTGGTATATTATG
ACAAGTGACATTAATCATAAAGATACAATAGAGTATTTTAAACAACATAAA
TATTTTAACTATGATGCCAATCATATTCATTTCTTTAAGCAAGATAACATT
GTTGCTTTAAGTGAAGAAGGAAAGCTTGTTTTAAATAGAGATGGACATATA
5 ATGGAAACACCTAATGGTAATGGGGGTGTATTCAAGTCTCTTAAGAAAGCA
GGATACCTTGATAAGATGCAACAAGATCACGTCAAATATATCTTCTTAAAT
AACATTGATAATGTCTTAGTTAAAGTTTGTAGACCCGTATTTGCGCGTTTT
ACAGTGACACAAAGTAAAGACATCACATCAAAAACAATTCAACCTAAAGAT
AGTGAAAGTGTAGGTCGGCTTGTAAATGTTGATTGTAAAGACACTGTGTTA
10 GAGTATTCGAATTAATGTTCATATTTCCAGGAGCTGCACTCCAGTTCATTT
CTGAAACTAAAATACTTCCATCGCTATTCACGCGCTCTACAAACGCTACGT
GACCATAGTAACCAGCGTCAGTTGTGCAATTGAGCCTACTGTAGGACGAT
AATCAATAGTATATCCATCAGCAGCTGATGCATTGTCCCAATTA

15

Sequence 3363

step.1002a08.cons.ok

TTACCTCCATTTAAGATTCAAAGTGATATGTCCAAAATGGCTCTATTTTTTA
AGAAACTTTGTGCTTGACCATCTATTTCTGGGACTAGCAAGTTGTTTTAAAA
20 ATGGTCTCTGTTGTGAAGCATGTGCTTCAAAAAGCTTAAAGTTTATATCAC
TATATTGAGATATGTTCATTTTGAATATCAGGTTCTCCAAGAATTTCTGATG
CATCATTTGCTAAACGCTACAAGTGAAGACGGGGTCGATCTGATTCATGCA
TGCGTCCAACGTGACGTACTACAGCTTCTGCAGTTGCCCTCGTGATCAGGGT
GAAC TGCAAAATTTAGGATAGAACGAAATAATTAATGATGGATTTGTTTCGT
25 CAATAAGTGATTGAATCATTTGATCCATTTGATCGTAAGGTTCAAATTCAA
CAGTTTTATCTCTTAACCCCATTTTCCTTAAATCTGTAATCCCAATGCTT
TGCAATGCTTCTTCTAACTCACGTTACGTATAAATGGTAAAGATTCTCTTG
TTGCAAAAAGGAGGGTTACCTAGATTACGTCCCATTTGTCTTAGGGTAAGAC
ATGCATATGTGACGGGAATACCTTTTCAATATAACTTGCATAGTTCCAG
30 CAGACGAAAAAGTTTTCATCATCAGGATGGGGGAAATCACAAGTACGTGTC
TTTCATCAGTCATGATTGGCAACCTCCCTTTATATAGTAAATGGCTGTTTAC
TTATTTCTATTGTAGCAGCCAATTGTCCTTCATAATTAACCTGCAATTA
AAAATTCGTTATTCTCATTAACCTCATAGTGCCTAAGTCTTGCACGTAAA
CCCAGCCACCATCTTTTAATTTAAGACCTACAGCATATGGATCTTTGTGCGC
35 CGCCTTTAAGTTGTGCATGTTCAAAAAGTCACGACAATATTCTTAAAAATG
TTCCAGCGTTAAATACGCGTTGATCGAAATGATTGTCATAAGCACCATTG
TTGTTTCAACGTTAGGTAAACAGGCTGATTGTAGTAAGAAGTTAATAAAT
CTAGCACTTCTTGTCTTTAATTGGTTCCAACACGTTCACTCCCTTAAACT
AAACAATATGTGAGTGACGGACTACCGATAAACATATTATTTACCAGCAA
40 CTCGTTTATCTATCTATTTTACTAGAAAGCTATGTATATTGTACACGAAT
GCTCGATTTAATTAGGAAAAATTCAGAAATAACGTTGAATATTACATATA
GAGCGTTAATATGTTTAAATCTTAATCATTTGTTACAGTATAATTAAAGA
GATATAAAAGAGTAAGGCGTTAGTCATCAATCATTTGTGATATTGAAATTAG
AGGTGTTTGATATTGAAAATTATTAATTTAGATTCAAAAAATCTTGCCTCA
45 TTTTATGTAGCATGCGAGTTGTTTAAACAAATACAGCAGCACCCCTCATGCC
AAACTCGGTTTAGCAACTGGTGGAACATATGACTGACGTATATCATTACTTA
GTAAATTTTATTAATCAAAAATAAAGTTGATGTGTGCGAGGTTGAAACATTT
AATTTAAATGAATATGTAGGTTTAAAGCAAGTCATCAACAAAGTTACCAT
ACCTACATGAATAAAGTGCTATTTGAGCAATATCCTCATTTTGTAAAAAT
50 CACATTCATATTCCTGATGGCTTATCAGAAAACTAGAAAGCAGAAAGCGGAA
CGATATAACATTTATTAGATGAAAGAGGGCCAAATCGATATTCAAAATTTTA
GGAATTGGAGAAAAATGGTCACATTGGTTTAAATGAACAGGGACTGACTTC
AATAGTGAACACATGTGGTGAACCTAACAGAAAGCACCATAAAAGCAAAT
AGTCGATTTTGTGACAATGAAAAGGATGTTCTTAGACAAGCAGTTTCAATG
55 GGGGTAAAAAGTATTTTAAAGCAAAAAGGATTATCCTACTCGCATTTGGT
CCAAAGAAAAAGAGGCTATAAGTAAACTGTTAAATGAACAGGTTACCGAA
GATGTACCTGCGACCATTTTACACACACACCCCTAATGTTGAAGTTTATGTA
GACGATGAAGCAGCGCCAGATTGTTTATAATTATATAGTTGTTTTATATAC
AAAATATTTTATATGTTTAAATAGTTTAGCGAAGGGTATACAACCTATAAGC

AAAAGAGAATTTCCCAAACGAAAGGATGAATTTTCATTGGAAC TACAATTA
 GCCATTGATTTATTAATAAAGAAGAAGCAGCAAAATTAGCTCAAAAAGTT
 GAAGAATATGTAGATATTGTTGAAATTGGTACGCCAATTGTAATTAATGAA
 GGGTTACCTGCAGTTCAACATTTAAATGAAAATATTAATAATGCTAAAAGTA
 5 TTAGCTGACTTGAAAATTATGGATGCAGCAGATTACGAAGTGAGCCAAGCA
 GTAAAATATGGTGCAGATATTGTTACAATTTTAGGTGTTGCTGAAGATGCT
 TCAATTAAGCAGCAGTTGAAGAAGCGCATAAACATGGAAAAGCATTGCTT
 GTTGATATGATAGCAGTGCAAAACTTAGAACACCGTGCTAAAGAACTAGAT
 GAGATGGGTGCAGACTATATCGCAGTTCATACAGGTTACGACTTACAAGCT
 10 GAAGGAAAATCTCCATTAGACAGCTTGCGTACAGTTAAATCTGTTATCAAA
 AACTCTAAGGTTGCAGTAGCAGGTGGTATTAAACCAGATACTATCAAAGAT
 ATTGTTGCTGAAGATCCAGATTTAGTTATTGTTGGTGGCGGTATTGCGAAT
 GCTGACGATCCTGTAGAAGCAGCAAAACAATGTAGAGCAGCTATTGAAGGT
 AAATAAGATGAGTGAATTTAATAATTATCGTCTTATTCTTGAAGAGTTAGA
 15 TTCTACTTTATCTCAAGTAGATAATACAGAGTATGAACGTTTTGCTAATGA
 TGTATAGGTGCAGATCGCATATTTACAGCTGGTAAAGGTCGTTACGTTT
 TGTGCTAATAGTTTTGCAATGCGCTTAAATCAATTAGGTAAAAATGCCTA
 CGTTGTAGGTGAGTCAACAACACCTTCAATTAAGAACATGATTTGTTTAT
 TATTATTTACAGTTTCAGGTTCTACAGAACATTTAAGATTATTAGCTGAAAA
 20 AGCACAATCTGAGGGTGCAAAAATTGTCTTATTAAC TACAAATGCGGAATC
 GCCAATCGGTAATCTTGACAGAGACGGTTGTTGAATGTCCTGCAGGTACTAA
 ACATGATGTTGAGGGTTCGAAACAACCACTTGGTAGTTTATTTGAACAGGC
 TTCACTTATATTCTTAGATAGTGTGTTATTACCTTTAATGGATGCATTTCA
 CATTAGTGAAAAACAATGCAAGAGAATCATGCTAATTTAGAATAACTAGA
 25 ATGAGAGATGAGCACTT

Sequence 3364

step.1002a09.cons.ok

30 AGTATCAGAATTGTCAATTAGATGTGTCGTCATTACCACATGCGCCTAAAAAT
 TAAAGCACTACTAAAAATTAAGACTAAAAACTTTTTCATATGTAATTCTCC
 TTTACTATATATCTTTATATTCAAACACTCGTAACGGCTCAAATTGAATAA
 CGTATTTGCTATACCGAGTGGAATAACCATATTTTGTGTTATAATGTTCAA
 TACTTTGTAGGACAAAACCTTCTGTAACTCAAAAAAATTAGCAAGTTTCAT
 35 ACAAGTCATGAATGCCTTGCAAAAATGCTTTTACAATATCTTTAAGAGGGA
 TTAACTTTTCATAGGCTAACCTACGTGCATAACCTTCAAATTTTCTATGAT
 TAAAAC'TACTTTGATCAACTATATTTCCATATGTAAGTTCGTTATGAGCGA
 GTTCCTCTGCAAGTGTTTCAAGTTTGAAGTAATAGGCAAGTTACGATTAA
 TTAATAATCATATCTCCAAGCCACAAACCAGATAACCTTTTAGGTAAGTTAT
 40 CACATTCAATGACTTCAATATAGTCATGTTCAATTAACATATCTTCATATT
 TCCCCACATAAAACACCCCTTATTTTCTTTTACTTCTTATGTATTCAGCGT
 AATCAAGAATTTCTTGCCATTCTTCGTCAAGTTAAATCACCATCAAGATGCG
 CTGCACGATGTTGTGGTTGTTCCCTCAACTTCTTCTATCCAACCCATCAAGA
 AAGCAGGGTTAACATTTAATGCAGTAGCTATACTTTCAATAGTATCATTTT
 45 TAAGATTTTAAATATCCCGCTTTCATAACGTTGTACAGTAGCTTCAGTTT
 TACCGATTTTTCTTCCCTAGTTCAGCTAAAGTCATACCTTGTTTTCTCTTG
 ATTGTTTCATTCTTTTTGAAAAGCACATCGTAATACAGCTCCTTTTACTTT
 ATAACATATATTATAAGGAAAAC'TTTCGTGATTGCAATATTTTCTTAAA
 AAC'TTTCGTAAAAATGCTTGACCTATTTTGTACAGCCATGATAAGATTACTT
 50 ACGTAATACGAAAGGTGGTGAATGAAATGCCTATCGACACTAAACTTTTA
 AAATCAAAAAATGGCTTTAAAAGGACACAACATCAAAACTCTTTCTGAAGAA
 ATAGGTGTTAATAGAGATACATTGTCCAATATGATACACGGAAGAATAAA
 CCTTCATATCCAGTAATAAATGGAATTTATTTGCGGTTAGATTTAACACCT
 CAAGAAGGAAGAGATATTTT'TTAAACAAAGACTTACGCAAAAAGAAAGTT
 55 TTAAC'TTAAAGGAGGAATTCAAATGACTGAAAACAAGGTATTATACAGGGT
 TTTTGTAAAAGAAACAAAGAATAGTGGATATACATTAGTTT'TTGAAAGTTT
 AGAAGAAGTCGAGCTAATATAAAAAATGCAGAAGAATTTGACGAAATTAT
 AATAAGCCCCCTCAAAAAGAGAGGCTCAAGATTAAACGATACTAAAGATAATC
 TTTTAAATGTTGTGAAC TGCAAATTACATTAGACCATGCGGCTTCACCAGT

5 CAGCTTACC AACGAACAAAC TATCATCGCTATCAATAAACATT TTTAATTTC
GTTTCTAATAGTTT TAGCATCATTAGTTGAATTAAATAAACAAAC TGATTT
ATTGATTTTAGCAACATTAGGATAATCCTCAATACGCTCTATTAAATTTGG
ATAATCCTTTGGTTATTTAAATCATACGCAATAATAAACTTTTCATATT
TTCACCTCCTTTTATAAGGAGTATAGCAGAAAGGAGCATAAACAATATGCA
10 AGATTTACAAGTATTTAATTTTGAAGATTTACCAGTAAGAAAAATAGAAGT
AGATGGAGAACCATATTTT TAGGTAAAGACGTGGCAGAAATATTAGGTTA
CACAAGATCTGATAATGCAATTAGAAATCATGTTGATGATGAAGATAAGCT
GACGCACCAAGTTAGTGCATCAGGTCAAAAACGAAACATGGTAATCATCAA
CGAATCTGGTTTATACAGCTTAATCTTTGACGCTGCTAAACAAAGTAAAAA
15 CGAAAGTATTAGAAAGAAAGCTAAACGTTTTAAACGTTGGGTAAACCGAAGA
TGTTTTACCTTTCCATTTCGTAAAACAGGTACTTATCAAGTTCCTGATAATCC
AATGGACGCATTGCAACTTATGTTTCGACGCACAAAAACAAACCAAGAAAGA
AATAGCAACTGTTAAAGCAGATGTTATTGATATCAAAGAAAATCAAAGCT
AGATGCAGGAGAATACGGATTGATAACAAAAACAGTTCATCAACGCGTTGC
20 TTATATCAGACAAAATTCACGGACTACCTAATAATAAAGAAGTTAACAAACC
TTTATATAGAGATATTAACAGTAACGTAAATACGATGGCTGGTATTAAAAAC
AAGAACAACAAATTAACAAACAAAAACATTTTCGATGACGTAATGAATATGATCAC
AAATTGGTTTCCATCTCAATCAACAATGTATGTCATCAACAATTAGAAAT
GGACTTTGAAAACGAAGTATAAGGAGTGATAGCAATGGAATACATTGGATT
25 TGCGGACGCTATCGAGTTTGTGAAAATAAGTGAATTTCTAAAAACGATTT
AGAAAAGCACGTTTATAGCAATAAAGAGTTCCAAAGAGAAATGTATGTACAG
ATTTGGCAAGAATCATAAACGCTACATCAAGATTAGACCGCAATTGACTT
TATAGAACAAAATTTAATGGTGTGAGAAACGTCACCTTAGAGGAGTTTAC
TGAAAGAAAGAGGATATATCGAAAAACCTTTTAGAAATGAAAGGAGAGAATGA
30 TAAATGAAATATCTATTAAAGTTACATGACGATGTTTATCGCAATGATCATC
ACATTACTTTTAGGAGGTGGTTTCTTTACGGTAATAGCATTTCATGTTA
ACACTTATCTTTAGCACATCTTCTGGGAAAAGTGGCTTGAGATAACAAAA
AAGACTGAACTTGCGCCAACAAGTAACAGTCAGAATCTAATCAAAATATA
CAACTTAATTTAATCAAAATATACGGAGGTAGTCAAGTTGAGACACAAATT
ATTA AAAACGCGCAATGATTTAAATA

Sequence 3365
35 step.1002a10.cons.ok
TCTATTTCTTTAAAAATGTTTGTCTTAAAGATTTTAGCGCACGTTT TTA
GTACGTACTTTTTCGTGTGGATGTATTTTACCATCTTCTTCAAACGCAAG
ACAGGTTTCATTTTCAATAATGTACCTACCCAAGCTTGAGCTCCAGTGATA
CGACCACTTTTGTAAATTTTAAATCATCAACAATTAAGTATGCACCA
40 ATATGTTGCTTATTTTCAGTTAGTTCATTAATAATATCATCAGGTTTATAT
CCCTTTTGTACCAATTGAGCAGCGTAAATTGCAAAGCTACCTTCAATCATC
GCAGCAAGACGGCTATCAAACGTATGTACTTGAATATCTTCAACCATTTC
CCAGCTTGTGTTGCTGAAGGATAGCTTCCGCTTATACCACTTGATAAGTTA
ATCAGCATGACATCAGTGTATCCTTGTTCTCTTAATCTCTCAAAATTTTCA
45 ATCCAATCGCCAATAGCAGGTTGGCTTGTGTTGGTATAGTTT TAGATGAA
GCCATTTT TTAATAAAATCATCTACAGAAAAATCATCACTCTCAGTGAAA
TTCACCTCCATCATCGAAAGTTACACTTAGTGAAGCGACTGGTATGTTATAT
TGTTCTATTTATATGTTGTGTTAAATAAATTGTAGAATCGGTCAATACTGCA
ATCTTCATCTTAATACTTATCCTCCTCAATTATACTCATAGTTATGATACA
50 ATAATTTT TAAAAAAGAAAATATCATCGTAATAATTAGAACCTATCTTG
TTACGTGTATAATAATGTTAGAAAACGAAATGTAAGTGGCAAAAATAAGA
TGGATAAATCCATAATTACTATTAAACAAGCACATTCAATTGAAAATGTGA
TAAGTAAATCACGCTTTATAGCATATATTAAGCCTGTTTCGACTGAAAATG
AAGCAAAAGCTTTTATAGATGAAATTAACAAAAACATAAAGATGCAACTC
55 ATAATTGTTTCAGCCTATACGTGCGGACCAGAGATGAATATTCAAAGGCAA
ACGACGATGGCGAACCAAGTGAACAGCTGGCATCCCAATGCTTGAAATAC
TGAAAAACAAGAGATACACAATGTTTGTGTCGTGACACGCTACTTCG
GTGGTATCAAGTTAGGTGCAGGCGGTCTTATTAGAGCATATAGCGGCGCG
TGCGTGATGTGATATATGATATAGGTAGAGTCGAACTAAGAGAAGCTATTTC

CAGTAACCGTTACGTTAGATTATGATCAGACAGGTAAATTTGAATATGAAC
 TTGCCCTACTACATTCCTTATTAAGAGAACAATTTTATACCGATAAAGTAA
 GTTATCAAATTGACGTAGTAAAAATGAATATGATGCTTTTATAGACTTTT
 TAAATCGAATTACTTCTGGAAATTATGATTTGAAACAAGAAGACCTTAAAC
 5 TATTACCTTTTGATATTGAAACCAATTAAATAAAAAACACTATGCAAGTTA
 AAGCAACGCGTGCTTATGACCTACATAGTGTTTTTTTATTATGAACGGTGA
 TGCTCATCATAATGATGTTGCTTACCATTTCCTTTTGTAAATTAATTT
 AATATTGGTCGATAATTATCATCTATTAATCCAGTAAATTCACGATCAAC
 TCAATCGTAAAGACAATGAGAATGAACATCATAAGTGCACCCAACGGTTGG
 10 GATAAATAGAGGATAACACTAGATAAACTAAACATAATCGCTATTGAATAA
 ATAAGTAAACGGTTTGTCTATGCGTATATCCTAAAGCAAGTAATTTATGA
 TGTAATGTGACTTGTCCGCTTGCAATTATATGTTGCCCTTTTTCATTCTGA
 CGAATCATTGCAATAATGTATCAATAAATGGCACCGCTAATATAACTATA
 GGAAAGAATAATGCAATAAATGTGATATTCTTAAAGCCGAGTAAGGATAAG
 15 AAACCGATAATAAATCCTATCATTAATGCACCACATACCTAGGAAAAATT
 TTCGCTGGGTGAAAGTTATAGAATAAGAAACCAAGTAAAGACCTTAAAGT
 ACACAGCAAAATCATGATAATAAATATGTTTCGCTTGTAATAAGCGATGAAT
 CCAATAGTCATTAATGCCAATGCTGAGACGCTGAGGCAAGTCCATCAAGT
 CCGTCGATAAGATTAAATAGCATTTGGTAATTGCTACAATCCATATTACTGTA
 20 ATAGGAATGCTGAATATGCCAAAATGAATCGTTGGACCAATTGGCAATGAA
 ATAAAGTCTATTGTAATTCCATAAAACGTAACAAATTAAAGCTGCAACAATT
 TGACCTGCTAACTTTAAATAAGGTCTTAGATCGTAAATATCATCAATCAAT
 CCAACCATATACATTGTAATTGCACCTAATATAAGCGGTTTAACTCACGT
 TCAATAGGGTGTCCGAGCCAAATCCCTATTAAAGAAAGAAAATAAAATGACC
 25 GTTCCTCCCATCACTGAGATAGGTTTCGTATGTACTTTTCTGAAATTAGGA
 CGATCTACTAAATCTAATTTTTTTTGATATTACAATAAATGGGTGTAATT
 ATTAACTGACTATCATAGTAAAGCTATAAGTAATAGTGATACATCAGT
 TCACCTTCATGTAATAAGTTGTACCTCACGCTACCACTGTTTTGTGATTG
 CTTTCACTTTTAGTTCTTTTCACTGAAGATTTTTTGACTATTTTCAACATCA
 30 TCTTTTCTTAAACATCATTTGCCTTTTCGGGATACCTCGCTTAAATCAACG
 CCATTTTTACGCGCATGTCCGACGTGCAGATGGTGTGTCATTGATACGTTGC
 TGGCTATTATCTTGTGTATTTTCACTTTCACTAGAAATTCGGATTGATTGT
 TTTGGCGTTTCTTCTCTTTAGCTTCATCTTTTTGTGAATTTTATTATCTT
 TAGCAACCTCTTTTAATTTCTTAGTATCTTCATGCTTAACATAATTATTTA
 35 ATTGCTTTTCACTTGAATGACACGCCCATGAATCAAAATAGAAAAAGTTA
 CAAAAATATAACAAATTGCTAAAAAAATTATTATACCAATAACATAGAATA
 CTTTTTTCATTAAAGTTCACCTAGTTTTCAAAAAAATTTTTATAATTTATG
 ATAACTTAAAGGTGATATTTTTTCTACATAATATATTAAATTTTCAAGGA
 GATGTAAAAAAGTTGAAAAATTTTCGCAAACTAATTTTGGTGGGCATTTTA
 40 GTATCGGGGTCAGGGATAGCGAGTGTACAAACAAATATAACTCACGCAAAA
 GAAAGTCACGATTCAACTCCTCAAATATTAAATTAGTGGGAACGTATGAT
 ACTTCTCAAGTTGATTCCAAAACGATGAAACAATTTAAAGAAATAGAAAAA
 GAAGATAATAATTTCCACATAACTAAAC

45

Sequence 3366

step.1002a12.cons.ok

GGGAGTGGGCATTTATGATTCAAATAAAAGGAGCAGTCGATTTCCCTATTT
 CATTGGATAGTACGACTTGGATATTTGACGATAGAAAAGTTACTATTGATG
 50 ATTAGAACGTGGGGTATTTGATGGTACTAGACCCATCAACTTTGATGATA
 ACTAAGGAATGGAACCGTGCTATTTTAGAAGGACAAACAAATCCACCAACGC
 TTGATTCTGAAATTAAATATAAAAAACGATCGGTTTAGATGAAACATTTG
 TGATAAATATGACACCATTTTCAAAAACGCCGAACCTATGGAAAATGCTA
 GCACAATAAACTTTCTAATAAAGATTATTCCATCAATGTACCCATGGACT
 55 TATTACCTTATTTGTTCTTTCAATTTTCTAGTAATGGCAAACGTTTATATA
 GTGATAATGCTGTTGATAGTTTCATATATATTCCTGAAGAAGGATACCTC
 ATCAATTTAAGTATGTACACATATAGAGGTGATTTAATATGAGAAAGGTT
 CAATGTATTATTTGCGATACTAAAGTTTTTATTGATGAACATACAGTCGAG
 GCTAAACGCTTAAAAAATAATCCTATACGCACCTTCATGTGCGATGATTGT

5 AAAAGTAGACTCGATACACCTAAACAAAGCAAACGAGATTAAAAATGATATA
TTGGAATTTCAAACCTATCTCACATAAAATAAGAGTGTGGGACTTAAAGTT
CCAAGATAAAATAAAAAAGACAATTTCTATTGAAATTATATAGAAATTGTCT
TTTTTATTATTTTTTTGATTATTTTCAGCTCGTTGAGCTGTTACTTTTCTT
10 ATATTAAGTGCCATTAGCACAAATCCTAGTTCTCTTTTGACTTTATTTATC
CCTCGAACGGACATTTCGAGTAAAACCCAAAATAGCCTTCATAAAATCCAAAA
ACAGGCTCCACATCAATTTTTCTTGACTGTAGATGC'TTTTGT'TTTTGGTT
CTGAAAAC'TTTTAAATAATTTGGACTTTAAAAATATCCCAATTATAATTTT
TCATTATTTT'TATTTTGCTTTTGAATCGATTAAATGTTATATAAGAAGGC
15 TTTTGATTTTATGATAGACACATCATTCGAATATTATCATTAAGTATTTT
TCTTACTACGCTCCTGAAAACACAGATTGGGTATAGGCATATAAAACTACTT
TTAACATTATTTTGGATGGTATGATATTGCATCACGATGATGTCTGAATT
CAGTCTCTAGAATTGTTTCTACAATATCATTTACATGTCGTGAAATATCAT
TTGTGGGGATAAGAAGTGAAGTTTCCATTGGTAGAGTAAGTTGAGTCATGT
20 TATAATCTTTATACATAAGGCACCTCGTTAATTTAGTTTAGTGATGTTTAT
TAAATTATACGAAAGGGCCTTATTTT'TTAAAGTATTTTAAATGTAATAATTA
CATATGAATACAAAGTATTTTGGCGAGACTCTTGAGGGAACAGGACAAGCT
GAAGACTACAGCTGAAAGCTGTCTCTAAGAAAGCGAGCCAACAATACGAA
GTATTGTAAATAAAGAAGCCAGTAAATGAATTTATGAAAAC'TCATTTACTG
25 GCTATTTTGTCTAGGAATTATGTCTCAGGCTCTTTCTACGATTAGTCAAGT
ACTTATAATTAATAAAGCATTACTTAACTAAACCGTTCATTTTAAATTTT
TTGAATAAGGTGTTAAACTATATTTGATTGTACGTTTGGCTAATGTGGTGA
CATTTGATAATCGTACAATTTATGATTCAATACAAGATAATGAATTGTTT
30 TAATAATCGATTATACAAGCAATGATGGCAGTCTTATGAGGTTTCTCAT
AGGCTGCTTTCTTAGTTTGTAGTAATAATCGACGACATGATTGTCATAATG
ATGCTGCCCTCTTATTATATTCATAATCACCCAAAATAAAAGTTTCTCGC
TTTTTTATTACCACGCTTGTGTGATGGTATCTCTACAGTGTGTATGACCTGA
TTGATATCGTTTGATATCAATGCCAACAAAGCATTGAGTTGTTTATTTGA
35 TTTAAATCGCTTAATATCACCAATCTCCCCAATAATCATAGCTGTGCTTAG
CTTACCAATACCAGGTATCGAATGAATATTTTCAAATAATCGAGTTGTTG
TGCTAATTGAATCATGGCATCATCTAATTGTTTGGAGTATGAATAGATTG
TTTTAATTGTTGAATAAGTAAGCGTAATTTTTCGACTAGAAAGGAATGTCT
ATCGACATTAGGATAGCTTTCTTGAGCAATCACCC'TTAATTGAAGTGCATA
TTTTGTAGCTTTATCCATTGACAT'TCCCTTATCTGTAGAATTGAATATATG
40 TGTAATCAGTACCTCCTTGTGCGATATCAAGAACCATGTCTGAATGAGTAAA
GATTTCTGCGATGTTGAGTGAATGATTGAATATCGACTACTAAACAATCT
TCTAAACCAGGGAATGTTTGAATGGAGTAATTCAAGGATCTGAAATTTAAG
TCGATTTTGTTCATTCTCGATTCTTAGATGAAAACGGACGCGTTCTCTTAA
TTCAAAGAATATTAACATCATGTATAGGTAAGCTGTCTGTTGTTAAGCGT
45 CGGTCCTAAACAAGCAAGCTTATGAGCATCTGCCGTGATCAGTTTTCATGA
TCTTAGAGCGCTCGTTTTTAAATTTGGCTTCTAACGGATTCATTTGAATATA
GTTAATTTGATTTACACAACAAAATCGTTCCATACCTCTTGAATAGATACC
TGATAGATTCAAAAATGAGTTGTGGGTGGTCTAAGTCATTCAAATACTTGAG
TAAATAATTGTAACCATTTTATTTATTCTGGATGAAAAACTCTTTTGGAA
50 TTTTCCATTTTATATAATGTGCAACTACACTACTTCTTTTACTAATATCAAC
ACCTAAGTAATCGATAAAAAAACCTCCTTTGAATAATTGAGAAGCTAAAAA
CTTTACTTAACTTTTCTCATTTTCATTTTCTTATACACGGTTTCAAGAACCC
AACATACTACAAACGAATTTCAAAGGCGAGAGTAAAGCTGACTTGT'TTT
TATACGGATTTAAAAATCCAAGAGTCTGGACAGTCTACTTCTCTATAACT
55 ATAAAAATAGCTATGAAAAATCTATCGTCATAGATTCTTTCATAGCTAA
TCTTAGATGTTTGTAAATATATTTAAATTAATTGTTCTGCTTGCATTAAATG
TTTAAC'TCTTTTGTCTCTTTTTCACGGGCAGATTTATTAAGTATTTTCTT
TCTTAAGCGAATACTTTTCAGGAGTACTTCCACCAATTCATCATCATTGAT
AAATTGTAACGCTTCTTCTA

Sequence 3367

step.1002b02.cons.ok

CTAAAAAATAAATTGTATATAACTTACTTTGAAATTGATCGAATATATAGT

ATCTTTAATAAAATGCAGATATTGTGGCGTAGAATTTGAGAATCAAAAAAT
GATTAAATATAGTTTGAAGAACTGAGCATAAGTACTAGAAAAATAGCCAGT
AAATGAGTTTATTATGAATTCATTTACTGGCTTCTTTAAATAATGATTTAG
ACAGTTTGCTTTTCTAGAACTACTTTGAAATGCGTTATTTCTTTAAGAAT
5 ATTAATATATAACTTGTGTTGCTAATGCTAACCTTGGTCAGAGTGTATTGT
CATGCGATAGTTAAGATGAGATCTTTTAATTAACCTTGTTTTAAACATC
GATGACAAGGTCTAATGTAGGAGGTGTAGAAATAATTTTCAGAATGATATA
AATAAATCCATAAATGAAGATAAATCGAACTATTTATACCGTAGGGGGT
ATATGTTATGGTATAAGTAACCTTACTATCACGTTTTCATTGGGAGGGGT
10 AATTTGAATAAATAATGGTGAAGAGCATAATCATCAAAATCACATGAATCAT
TCCAATCACATGCATCATGATAACCATGCCTCACATCATCATAGTGGCCAT
GCACATCATCATGGAATTTTAAAGTTAAGTTTGTGTTTCATTAAATTTT
GCAATACCTATCATTTCTCTTATCGCCAATGATGGGTGTTAACTTACCTTTT
CAATTCACATTTCCAGGTCTGAATGGGTAGTGTAAATATTAAGTACAATT
15 TTATTCCTTTATGGTGGTAAACCGTCTTGTCTGGTGGTAAAGATGAAATT
GCTACAAAAAACCAGGCATGATGACCTTAGTTGCTCTAGGTATTTCAGTA
GCTTATATTTATAGCTTGTATGCTTTTTATATGAATAACTTTAGTAGTGCA
ACTGGTCATACAATGGACTTTTTTGGGAATTAGCAACCTTAATTTAATT
ATGCTATTAGGACATTGGATAGAAATGAATGCTGTCGGAATGCTGGAGAT
20 GCTTTAAAGAAAATGGCAGAACTATTACCTAATAGTGCTATTAAAGTTATG
GATAATGGCCAACGCGAAGAAGTTAAAATATCAGACATCATGACTGATGAT
ATCGTCGAAGTAAAAGCCGGAGAAAGCATTCCAACAGATGGTATTATCGTT
CAAGGACAAACATCTATAGATGAATCCCTAGTCACTGGAGAATCTAAAAAA
GTACAAAAAATCAAAATGACAACGTCATCGGGGTCTTATTAATGGGTCT
25 GGAACAATACAAGTCAAGTTACAGCTGTTGGAGAAGATGGATATCTTTCT
CAAGTTATGGGACTTGTAAATCAAGCACAAAATGATAAATCTAGTGCTGAA
TTGTTATCTGATAAAGTAGCGGGTATTTATCTACTTTGCTGTAAGTGTT
GGCGTGATTTCTTTTATTGTCTGGATGCTCATTCAAAATGATGTTGATTTT
GCATTAGAACGTCTTGTAACTGTGTTAGTCATTGCTTGTCCACATGCTTTA
30 GGCTTGSCAATACCTTTAGTCACTGCACGTTCTACTTCAATTGGTGACAT
AATGGTTTAATTATTAAAAATAGAGAGTCTGTAGAAATAGCTCAACATATC
GATTATGTAATGATGGATAAACTGGTACTTTAACTGAGGGTAACCTTTTCT
GTGAATCATTATGAGAGCTTTAAAAATGATTTGAGTAATGATACAATATTA
AGCCTTTTCGCCCTCATTAGAAAAGTCAATCTAATCACCCATTAGCTATAAGT
35 ATTGTTGATTTTGCGAAAAGTAAAAATGTTTCATTTACTAATCCACAAGAC
GTTAATAATATTCCAGGTGTCGGATTAGAAGGTCTAATTGATAATAAAACA
TATAAATAACAAATGTCCTTATCTTGATAAACATAAACTTAATTATGAC
GATGACTTGTTTACTAAATTAGCTCAACAAGGTAATTCAATCAGTTATTTA
ATTGAGGATCAACAAGTCATTGGCATGATTGCTCAAGGAGATCAAAATAAA
40 GAAAGCTCAAAACAAATGATAGCTGATTTACTATCAAGAAATATTACACCA
GTCATGCTTACAGGTGACAATAATGAAGTGGCACACGCTGTCGCAAAAGAA
TTAGGTATTAGTGATGTTACGCACAACCTCATGCCAGAAGATAAGGAAAGC
ATTATAAAAGATTATCAAAGTGACGGTAATAAAGTCATGATGGTCCGAGAC
GGTATCAACGATGCGCCGAGTCTTATAAGAGCCGATATTGGTATAGCAATT
45 GGTGACGGCACAGATGTTGAGTGGATTCAAGTGATATCATACTTGTTAAA
AGTAATCCATCAGATATCATTCATTTCTTGACTCTTTCAAATAATACTATG
AGAAAAATGGTGCAAACTTATGGTGGGTGCAGGTATAATATTGTTGCT
GTACCTTTAGCAGCTGGCGCATTAGCTTTTATCGGGTAAATATTATCACCA
GCTGTAGGAGCAATATTAAATGCTTTAAGTACAGTTATAGTAGCGATTAAT
50 GCTTTTACATTTAAATTTAAATAAAAAGAGGTAACCTTATGTATAATAAAG
TTTTTGCAATTTTAATTATAATTTTTCCATAATAATTATTGCGTCTAATG
ATACTTTCGAGAAAGTAAGAATGATATGATGAATATGAAAGAAGATAAGA
AAAAATCAATGGATATGACAAATATGAAACATCATGACGAAAGAAAGAAAT
TAAATTCTTCACAAGGAAAAAATGAAATAATATTCTTAAAGTTGCAGAGT
55 CAAAAAAGATAACAATGGTTATAAAAAATTATACATTTAAAGCTCAGGAAG
GAAAGACAGAGTTTTACAAAAATAATTTTTCTAATACTCTAGGCTACAATG
GAAATTTACTTGGACCAACTTTAAATTTAAAAAAGGAGATAAAGTTAAAA
TTAAGTTAATAAATAACTTAGATGAAAATACAACATTTTCATTGGCATGGAT
TAGAAGTAAATGAAAAAGTGGATGGAGGGCTTCTCAGGTTATAAAACCAG

GAAAAGAAAAAACTATAAAATTTGAGGTAAATCAAGATTCTGCTACGTTAT
 GGTATCACCCCCACCCCTCTCCAAATACAGCTAAACAAGTTTATAATGGCT
 TATCAGGATTATTATATATAGAAGATAGTAAAAAGAATAATTATCCTAGTG
 ATTATGGAAAAAATGATTTGCCATATAAATCCAAGATAAAACATTTGTAT
 5 CTAaaaaaTTAAATTATTCAAAAACGAAAGACGAAGATGGCACTCAAGGTG
 ATACTGTTCTTGTGAACGGAATAGTAAACCCCAACTGACAACAAAAGAAG
 AGAAAATACGTTTGAGACTTTTAAATGGTTCTAATGCTCGAGATTTAAATC
 TTAAGCTAAGTAATAATCAAAGTTTGGAGTATATTGCTTCAGATGGCGGTC
 AATTAAAAAACGCTAAAAAATTAAAAGAAATTAATTTAGCTCCTTCAGAAA
 10 GAAAAGAAATAGTAATAGATTTATCTAAATGAAAGGCGAGAAAATCAGTC
 TGGTTGATAATGATAAAACTGTAATTTTACCGATTAGTAACAAAGAGAAAA
 GTTCTAACAAAGGTAATACACCAAAAGTAAGTAAAAAATAAAATAGAAG
 GTATGAATGATAATGTTACCATTAAATGGTAATAAATTCGATCCTAAAAGAA
 TAGATTTTACACAAAAGTTAAACCAGAAAGAAGTATGGGAAATTGAAAACG
 15 TCAAAGATAAAATGGGTGGTATGAAACATCCTTTCCACATCCAT

Sequence 3368

step.1002b04.cons.ok

20 TGATAAAATTAGTGGTTAGCTATATTTTTTACTTTGCAACAGAACCGTTTTT
 TGTATATAGTCTTATAGGAATTGCTATATTTTTTGTGCCGGTAACGATTAA
 TAATAACTCTTCAATTTTATTAGATCATTTTGTACATGGATAAGTAGCGC
 ATTACCTCTTTTAACTAAGATATTCATAATGATTATCATTATACCTAGGTGC
 TATTTATCCATTTATTAAAGGGACATGGAATCGGAATACCGTTGAAAACAAT
 25 TTTTAGTTTATTTAAAGTTTGGGAGTCATTATAGGCGTTTGTAAATTTT
 TAACATTGGGCCAAGTTGGTTACTTAAATGAACAAACGGGAATGTATGTTTT
 TAACATTTGGTAAATCCGGTAGGATTAACAGTACCTGCAGGAGGCGCGGT
 ATTAGCTTTATTAGTAGGATATGGCTTATTAGAATTTGTAGGTGTTTATGC
 GCAAAAAATTATGTACCCGATATGGAAAACGCCTGGACGTTTACAGCTTAA
 30 TGCTTTAGCATCTTTTGTGTAGTTTGTGTGGGTTTACTTAAACGAA
 TAAAGAGTATAAAGAAGGTAAATTCACGGAAAAACAAGCTGTTATCATAGC
 AACCGGCTTTTCTACAGTTACTGTAGCTTTTATGATAGTTATTGCTAAAAC
 CTTACACTTAATGGATATATGGAATTTATATTTTGGTCTACCTTGTTTGT
 TACTGCTGCAGTAACAGCTTGTACAGTTAGGATTTGGCCTATCAGTAAAAAT
 35 TAGCAACACATATTATGATCAGCCATTTATAGAAGAAGATACAAGCGAATT
 AAAAGGTTAAAAAAGCTACTTTTGTGCATGGGAAAAAGCAATGGAACTGT
 TGAATCATCACTAACGTGATGAAAAATATTTGGTTGAATTTCAAAGAAAG
 TCTGATTATGACTATGAATATCTTACCCACCATATTATCAATAGGTTTAAT
 TTGCTTGTACTCGCAGAATATACAGTGATTTTCGATTATTTAGCATATGT
 40 TTTTATCCATTAACTTGGTACTTCAAATACCAGATTCCCTTTTAACTGC
 AAAAGGCGCAGCTATTGGTATAACAGAAATGTTTTGCCTTCATTAATTGT
 AGTCGAAGCACCATTAACTACTAAATTTATAATTGCTGTTACTTCTGTTTC
 TACAATTATATTCTTTTCAGCTAGTGTGCCTAGTATTCTCTCTACTGATAT
 ACCCATCCGCATAAGAGATTTAGTGGTTATATGGTTTGAGAGAAGTGTATT
 45 GAGTTTAATTATAGTAACACCTATCGCATATATTTTTTATAATTCAATAA
 TAAATATCGTATATAAAAAAATCACCTTTTTTACAAATATTTATAGTGAT
 AAGGAAAGCAAGGATATGAAATTTCCATCTTAATTTACTGTTTATTTT
 AAATCAATTATATATTTAGGGTACACCCTAATCATACAAAATGAATTGTT
 GAAATTACAGCAATTAAAAATGATGATTAGGGTGTATATTAATTATGAGGT
 50 GAACTAAATGATTGTTGGGCTATGCGCGGGTATCATCAGTAGATCAAAAT
 TAGAAAAGACAACCTGAAAAATTAAAAACATTTGGTGCTGAAAAAATTTTA
 CAGAAAAACAATCAGGTCAATCTATAGAAAACAGACCCGCTTTTCAAGAAG
 CATTGAACTTTGTGAGAATGGGAGATCGTTTTGTCGTGGAATCAATTGATC
 GTTTAGGTCGTAATTATGATGAAGTCATTAATACCGTTAATTATCTTAAAG
 55 AGAAAGAAGTACAATTGATGATTACCAGTCTACCAATGATGAATGAAGTCA
 TTGGTAATCCATTATTAGATAAAATTTATGAAAGATTTAATTATACAGATAT
 TAGCAATGGTTTCAGAACAAAGAAAGAAATGAAAGTAAACGTCGACAAGCTC
 AAGGGATTCAAGTTGCGAAAAGAGAAAGGCGTATATAAAGGACGACCTTTAT
 TATATCTCCCAATGCGAAAGATCCTCAAAAACGTGTTATCTATCATCGTG

TTGTAGAAATGTTAGAGGAAGGCAAAGCGATTAGTAAATTCGGAAGAAG
 TGAATATTACAAGACAACTGTTTATAGAATTAAACACGATAACGATGATT
 TGGGTTCTGTTGCAAAGTAAAAAATATAGCTAACCACTAATTTATCATGT
 CAGTGTTTCGCTTAACTTGCTAGCATGATGCTAATTTTCGTGGCATGGCGAAA
 5 ATCCGTAGATCTGAAGAGACCTGCGGTTCTTTTTATATAGAGCGTAAATAC
 ATTC AATACCTTTTAACTTCGAAATGTTTTTTTATTTCTTCGGCAAAATAT
 TGGGGTGTTAAATATTTCGGAGGCATTTGACTTAGTTTCACGAGCATAATTT
 ATAGATTACCCCAATTGTTGGCCTTCGTTAATCTTCGTTTCTATATTATGC
 TCTGTCGTTTGTATATAGACTCTACCTTGATATGGAGGCTTTTTATCCGAT
 10 TTATAATGGTCAAATTCAAAGATAGATTGTGCACTTTGTAAACCTAAAGTT
 TGACAAATTTCAACAATGTTACCATGTTTAGATTTGAAAGTATCTAACAAT
 AATTCACAATCCGTGACATGAACGCTTTTAAATATTGAAACAACCTGACCC
 CAAACTTTTAAATAACGTTGATATGATAAGATTTTTAAGTTTCCTAGTCCT
 ACAGTTAATAATCTTTTAGGCTTTTCATCAATGTAGATAAGCGTTGATGAT
 15 ATTTTGGCGATTGAGCTATTGATGACATGATATTGTTTTAATGTTTCTAAT
 CGTTCCTTTAATTAACGCTCTTATGATATTTTATCTCACCTAATTGGTTGAGG
 TGATCTGGCACACCCACCACCAATCGTTTCAATTGATGTTGGCGTATCTTGA
 TTAATAAAAAAATTGCATATAGATCCTCCGTTCCCTCATCTGTATAACAATA
 ATTATAAGTTATGAACACCCCAATATTAACAACCACTTTTACTCTAA
 20 CTCTAAATAAAAACCACTTTTGTTTAATTAAGTGTGAAATCTTCTTCATC
 ATATTTTCTGTATAAAAAATAACCTGGGACAATCATTTTGTCCCAGGTATT
 ATTAATAATATAACAATCGATGACAAAATTAATTTGAGTTGTTTACATATT
 TTAACCTCTAGTCATCTTTGCTAAGGGAGACAACAAATTATGTCCTCGTACCT
 TTTTATTGTTTATTTATATCTATTACTTGATTAAAAATTTTCCTTTTAAAA
 25 GGCTAATCCTATGCCACCAAGTTTGTAATCGCACGAGTATCAATAACTTT
 CTTCAAGAATGCTGCTTTTTTACCAGCGAT

Sequence 3369

30 step.1002b05.cons.ok
 CAATATGTTTCGAGTAATGACTTGCGCGCATCGTCTTGCTGTTTCAAATATT
 CAAATAAAAAGCCAGTCCATACCCAAATGTTTTTGTATTAAGAGCACTAC
 AAGGCGTATTGCTGCATCATTTTCTCCTGTAGTCATGACTTATACAACGTG
 AAACCTTATCCACGTGTGAAACGTGTAACAGCAATTAGTTTCATCAGTACAA
 35 GTTTTATGTTATCAGGTGTGTTAGGGCAAAATATGAGTGAATTAGTCGTAA
 GTTATTTGAATTGGCAGTGGGTATATTTTATTTTAAACAATCTTATATTTAA
 TTCTCGTATTAGTTATTTTATAAAAAATGTACCTGAGAGTCCACATAAAAATC
 CTGATATTCAACTCATTAAGTTTTTCAATAACTTTAAAGATTTCAAAGACA
 ATCTTAAAGTTTTCTATTGTTTATTTATTTTCACTAACATTACTGATTATGT
 40 TTATAAGTATGTATGATATTTTAAATGAATATGTCACATCACACCAAGTTG
 GGGGAGACATGTCGTGTCCTCAATGATGAAATGTTTCGGTGTGATAGGCA
 TAGAACTAAAGTAATCATTTTATTTTAAAGTAGATTTTAAATTTTATAT
 ACATATTAGCGTTGTGGAATAGGTGGAAGGGGGAGTCAAAAGGTGAAAAC
 GCTAATGATTAAAGCAATGGGGACGGTGATACGTTTATCGATTGAGCATCA
 45 ACATCCGGATACATTACTTCAAGAAGCTGAAATAAAAAATTCGTGCTTGGGA
 ATCACAATTTAGTGCTAATGATCCGAAATCAGATTTGATGAATGTGAATCA
 GCATGCAGGTATCGCACCAAGTCAAGGTTAGTTCTGAGATGTTTAAACATGAT
 ACGTTTTGGTTACGAACTACATTATCTTCTAATTTTAAAGATGAACATTTT
 GATAGGGCCACTAGTCAAAATATGGAATAATGGTTTTTAAAGATGCATGAA
 50 ACCTAAAGAAGAGGTATACAACGTGCTTTATTGTGTATGGATCCTGAAAA
 TCTTGTCTAAATTCAAAAACACATGAAGTATTTCTTACACAATCAGGAAT
 GGAGATTGATTTAGGAGCTATAGTTAAAGGCTATTTTGCTGATCAATTACA
 GCAATACTTTTTAGCTCATGGTGTATCTTCTGGCATTATCGATTTAGGTGG
 TAATGTTTAAACAATTGGTAGACAACCCGAAACATTAGAAAAATGGCATGT
 55 AGGTGTACGTAATCCATTTTATAAGGATACACTACCACTCGTTACATTAAG
 CGTAGAGCATCAATCAGTTGTACATCAGGTATCTACGAACGCTACTTCAT
 ACAGGAAAAATCAATTATTTTCATCATATATTGGATTCAACAACAGGTTATCC
 TGTAGATAATGATATCGTAGCGTGACAATCATATCTGATCATGGGATTGA
 TGGCGAGGTATGGAGTACAATTTGTAGTTTGGTTCAGTCACAAAAAATAT

TGAATTATTAAATCTCATTGACGGTATTGAAGGCATTATTGTGACAAGAGA
 TGGAAGCGTTTAAATGACTTCGAAAATGCAAAAGTATTTATAAAAAATTTAA
 AGAATACTCATGAAGAGTATTTTATTTTAAAGTAATAATGTGAAAATATTC
 ACAAAAAAACTAGGAGGATTTGCTATGAGTAAGGAAATATTCGATACTTT
 5 AAATTTAAATGTGGTGCCGAATTAAAAAATAGAGTATTAATGGCACCCATG
 ACTATCCAAGCTGGGTATTTTGATGGAAGTGTTACATCAGAAATGATTGAT
 TATTATCAATTTAGAGCTGGTGATGCTTCAGCAATCATGTGTGAAAGTTGT
 TTTGTTGAAAATCACGGACGAGGATTTCCGGGAGCTATAGGTATTGATAAT
 GATGACAAAATACCTGGACTCAAACGTTTAGCAGAAGCGATTCAAGCTAAG
 10 GGATCAAAAGCGATTTTGCAACTTTATCATGCCGGAAGAATGGCAAAATCCT
 AAATTTAAATCGAAGGAGCAGCCGATATCTGCGAGCCCCATTGCAGCATTA
 AGACCTGATGCTGTACCACCTAGAGAAATGACACATGCTCAAATCAATCAG
 ATGATTGATGACTTTGGAGAGGCTACACGTCGCGCTATAGAAGCGGGGTTT
 GATGGTGTGCAAAATTCATGGCGCAACACATACCTTATTACAACAATTTTC
 15 TCTCCACATTTCTAATCGGAGACAAGATTCATGGGGAGGCAGTCGTGAAAAA
 CGTACACGATTTCCAATCGAAGTTTGAACAAAGTTCAACACGTCGTTGCT
 GAAAAAGAGGCTTCTCATTTTATTATAGGATATCGATTCTCACCTGAAGAA
 ATTGAAGAACCAGGAGCAGTACGTTTGAAGATACCATGTTTTACTAAATACA
 TTAGCAGAATATGAACCTGATTACTTCCATATATCAGCAAAACAGTTATCAA
 20 CGTACATCTATTGTGAATCAAGAAGATACAGAACCTTTAATTAATAAGTAC
 ATCAAAATGCAAAAGTGACAGTTGGCAAAAATTCATTAATTGGTGTAGGT
 AGTATTGCCCAACGACAAGATGCAGAACATGCCCTTGAAGTAGGATATGAT
 CTTTTAAGTGTGGGAAAGCCTATTTAGTGGAACCACAATGGACAGATAAA
 ATTTACAAAAACGAAGAAGTAGAACAATTTGTCGATATACATGATCAGAAA
 25 GTACTTCACATACCATCCCCCTTATGGAAAGTAATGGACTTTATGATTTTA
 GATAAAGAAGAAGAGCATCGTAAATATGAAAAATTTAAAGCACTTCAAAT
 AAAAAAGTTAAATTTAACAAGGTACGTATCATGTCTATGCAAAAGGTCAT
 AATGGCAACTTACCTATGAAAGTCCAATTATCAGAAGATAAGATTGTAAGT
 ATCGAGGTAGATGATAGCGGAGAGTCTGAAGGCATAGCGAACCAGTGTTT
 30 GAACGTTTACCTCAAGATATTATCAATGGGCAACACTGAATGTAGATGTC
 ATTTACAGGTGCGACAGTAACAAGTGAAGGCATCGTGAAGGTATTGCAGAT
 GCAATTGAACAAGCAGGAGAAGACCCAGATATTTTACGGGCGCGTCCATAA
 CCAGTCGTTTCAAGTGGTCTGATGAGGTTGTTGAAGAGA

35

Sequence 3370

step.1002b07.cons.ok

GATAAGAAATCTTAAATCAACTTCGGCAATATTATAGTCTATAATATACAA
 GTCAATTTATTAGTTGTTTCAATCAGACCGTATTTGCCATCTTTTCGTCT
 40 ATAAACAATGCTTGTGCCATCAGTTTCTCTATCGGTAAAAATAAGAAATC
 ATGTCCCAATAAATTCATTTGCAATACTGCTTCTTCAGAATCCATTGGTTT
 TAAGCTGAATTTGTTTGAACGAATAATTTCAATATCATTTTCACTTTCTAT
 ATCATCTGCATGATGTTTGTGTAGACTCTTGTACTTCAGCAACAAAGAT
 ATCTTGATCTCCAGCATCTCTATGTTTACGATTTACAGAGTTTGTATTT
 45 ACGTACTTGTCTTTCCAATTTGTTAGTAATTAATCAATGCCTGCATATAA
 ATCATCATGTCTTTCTTCAGCTCTAAGAGTGACATCTTTTAAAGGAATTGT
 AACTTCAATTTTAGTTGTAGAATTAGAATAAGTTTAACTCTAACATGTGC
 TACAGCATTTGGCACATTGTTAAAGTATCTTTCTAATTTACCTACTTTCTC
 CTCAATATAGTTGCGAATTGCATCTGTGATAGTGAGGTTATCTCCATGAAT
 50 TTCAAATCTAATCATAGTAATCTCTCTTAAACCTCTATATTGATACTTCT
 TACCTATATTTTACCATGTTTATGTTATCGTGCAACGCAACACTTTGAA
 TTTTCTGACTTTTATATCGTACAATTTACACCCGCGTGATGAATAGTTAA
 CCCTGTTGTATAAATATCATCTATGAGTAGTATTACTTTCCCTTCGATATT
 TATCTCTTCACTTTTATATAAAATGGATTAGGGGCTTTTGAACGTTCAAT
 55 CTTTCCCTAATTTGACTGCTTAGGACGTATATGTGTACCTAATACATCTTG
 ATATGAGATGCCCATTTTATCTAAGACAGTGGTCACAGGATTAATGTACG
 TTCAATGTCCGTTCAATTGGAGAAGGAATGGGAACATATAATCATATTG
 CGTTTGAGGTAAACTAATTTCTCGCCAATACCTCTGCCAAATAGAAGTC
 TCGCTTAATCTTATACTGGTGATGATTTCTTTTCATCATACCATCATAACT

ATAATCACAAAACAATTGTTCCATTAAATAAAATGTTTACTTAAGAATTG
ACAGTCATGACATTCGCCATCAGTAGAAATTTTGGTTGTTAAGGCACCTAGG
ACACCTACCATCTTTATGAATTTTGGAGTTAAGCCATAATGTTTCACATCG
TTCACAAAATATCTGAGGTGGTTTAAATAAGTTATAAATATGAATGGTCTC
5 ATGAATAGGTTGATTGCAC TGGATACATTTACGCATCAATCCATCCCCTTT
TAATTGCTAAACGATTCAATTGAAATAATGTTTCTTTTAGCTAAAATCATCG
ATAATGTAACACCTTCATGAAGAAATAAACTAAGCCACTTGGAGACTGCT
GTTTACGTCCTACGCGACCAGCAATTTGAATTAAAGCCTCTTGTGAAAAC
TTCAGCATCAACTACAACGACATCTAAGTGTGTCATTGTAAATCCTCTTT
10 CTAAAATAGTTGTAGTGAATACAATTTTGTGTTGTCCTCGTCTTAAAGCTT
CAATTTTTCAAAATCGTAAATCATCTTCACTGTGAACGCAAAATCAAATCAG
GGATGTCCATTTTATACTGTTGATACATTTTATTCATAATTTCTATATTAT
TAATAAAGACCAAAGTAAACGTTGTTGGTTAATTGATATCTAAATATAT
TAAGTAATAAATTTTGTTCGTTGATTTTAATTGAAATATTTGAACT
15 TAGGAATAGGAAGGGGGGATCGGTGAAAACGGGCTGGTAACCTAATTATTT
TTTCTGGGGGAATTGTTTAAAAAATGACGCGGTGGTGTGGCCGTCATGA
AAATATGTGAATGATTCGATTTTGAAGCAAGTTGATTGCATTTGATAATT
GTGGATCCATAGAACGGAAGCAATCTACCTCATCGACAAATACAGTAT
CAAAATGCTGTTTAAACCTCAATAATTGATGGATAGTAGCAATAACAAAAT
20 GACCATTATATTGCTGTCTACTAGATTGATGTAGCACATCTATATGTTTAT
CGATAAAAGCATCTTTAATTCGATGACTTATCTCAATAATAACGTCCTACAC
GAGGTGATACAATAGCAATATTTATGTCCCATCTGACGAGCTATCCTAATTC
CCTCAAACATCATTTTCAAGTTTTCCTGCTCCTGTGACGGCATATAACAAGA
GGTCATTATATTTTTTATAGCTTGAACGATAGTTTCTGAAGCGTATTGCT
25 GTTGTCTTTGATAATTCAAAGGGTAACTGATAATTCGCTTTTGTGCTTTTT
GAAAACTTTTTACGAGACATATATCTGTAATACTATCCATTCTGCCTAATT
GAATGCATCGTCTACAGTATGTGATTGTTTATCCAATATCGTTGAATTGT
AAGTCACAAAATATTGCAAGGAAGTGTTCACATTTGAATACAATACCAAC
GCTCAGTGTCTTTTATAACACCTTTTGAACCTTGAGATACTGATTTCATCTT
30 CAAGTTGATGCGTTGGAGAAATCAGTCTACCGTAATATTTAATTTTACTTC
ACTACTTTCAAAAAAGTCTGAAATACTCAAATAAGCGAACACAAGTGATAT
CTAATGATATTACCACTGTGTTCTTTCCCTATCAAGCTTTTCAAGCTTTTGTGAT
TATATATTAAATTAATGTGATTCCTTCTTGGGAAGTAACCCAATCCTAAACC
GCCTGATCCTAAATGTGATGCTATCACTGGTCCAAATTCACAATACGAAT
35 ATGAACATTAGGATGATCTTCCCTTAATTGCTGAAGAAATGACTTTCCATC
TTCAGTTTATACACGTTTATTACAAATACTGTACATCTTCCATGCCTTC
TATTTCTTTAAAAATGTTTGTTCATAAGATTTTAGCGCACGTTTTTTAGT
ACGTACTTTTTCGTTGGATGTATTTTACCATCTTCTTCAAAACGCAAGAC
AGGTTTCAATTTCAATAATGTACCTACCCAAGCTTGAGCTCCAGTGATACG
40 ACCACTTTTTTGTAATTTTTTAAATCATCAACAATTAAGTATGCACCAAT
ATGTTGCTTTATTTTCAAGTTAGTTTCAATAATATCATCAGGTTTATATCC
CTTTTGTACCAATTGAGCAGCGTAAATTGCAAAGCTACCTTCAATCATCGC
AGCAAGACGGCTATCAAACGTATGTACTTGAATATCTTCAACCATTTCAACC
AGCTTGTGTTGCTGAAGGATAGCTTCCGCTTAT
45

Sequence 3371

step.1002b08.cons.ok

ATGTACATGTTGGTGTGTTTGACGTTTCGATATTTTGTAACTTTTGGTAAAC
50 AAGCTTCCATGTTCCACATGCTTCTTCTGGGGTTTGATAATCAGAGGTTAG
TTGATTCAATAACATTTAATAATTGCCAGTATCTTCTGGTTCATTAATTAA
ATCTTCTAATATGATTTTAAATGTTTGTACATCTTGTCTTGAATATATTG
AATGAGTTTAGGAAAATCTTCCATCTTCTTTACGTATATGATCCAACAT
ACTATACGAAACTCGCGATATAAGTCTTGTAAATTTAATAAGTATGGGTG
55 ACTAGGACCATGTACTTTTGCCAAATTTAGTCATGTAAAGGTGAAGATTCTT
AAATTCCTTCTTAAAGCGTTTTCGTGATAAGCTGATTGTATATATTGTATAAG
AGATTCAACATTTAAAAATTTAGGATTAATGGTACTGTACCTTCTGTATT
ATCAATATGATTCAACTTATTTAATAAGGAATTTAAGTCAATATTTGGTTT
ATGATTGACAGCTGAAGCGATACCTCTTGTCCGCCACAACAAAATTCAT

TACGAAGATGTTGAAAAACGTAGGGAAGAATTTTCATCCTTCAGGAAGTCGT
 TCTAATTTAACCATTAAAGCAATGGGAAAAGAAATTCCTTATATCGAT
 GCGACATTTCCAATTATAAAGTAGGTAATACAAGGTTACTTATTGGAGAT
 AGTTTAGCAGGAAGTATCGCTTTAATGACTGCAATGACTTACCCAACATTT
 5 TTTAGTCGAGTTGCGTTATTGAGCCCAATGTATAATGAAAATATTAAGAAA
 AAAATTGATACATGTATGAATAAAGGTCAATTGACGATATGGCATGCCATT
 GGTTTAGAGAAGCAGATTTTATTTTACCAACTAATGGTAAAAGAGCTAAC
 TTTTAAACACCTAACCGTGAATTAATCAACTGATTAAGAAGATAATATT
 GAATATTTCTATAAAGAATTTAACGGTGGACATCATTTGGAATCATGGAAA
 10 CCATTGCTAGGAGATATTCCTTACAATTTTATAGGTGATCCAATAAATGGA
 AAATATGTTTAATAAAGTAAGATAAAATAAATAGTATGTTAGCGAATGGTT
 GTTGAATGCGAAACACTATTAAATTCAGAAATGTAATTGTTTTCATGATAA
 AAGTAACGGTTTTCAAAAAGTTTTGATATAATAGGAATAAGTTAAACAAAG
 GAGGAATTTAAATGATTTTAGGATTAGCATTGGTCCGTCAAAGTCATTTC
 15 AAGATGAGGTGAATGCTTATCGCAAGCGATATGACAATCATTATGCTCAA
 TAATGCCCTCATATACGATTAAACCTCAATTTGAAATCGATGATCATGATT
 TTAATTTAATTAATAAAGTGAAGTGAATAATCGAATTTCTAGTATTAACCAG
 TAGAAGTACATGCTACAAAGGCATCTAATTTGCTCCAATCAGTAATGTTA
 TATACTTCAAAGTTGCTAAAACAGAGTCATTAGATCAATTATTTAATCAAT
 20 TTAATACAGAAGATTTTACGGTACAGCTGAACATCCTTTTGTACCACATT
 TTACAATTGCCCAAGGCTTAACAAGTCAAGAATTTGAAGATATATATGGTC
 AAGTAAATTAGCAGGGGTAGACCATAGAGAAATAATTGAAGAACATATCGT
 TACTTCAATATAGTGAAGAAGAGGACAAATGGACTATTATTGAACTTTTA
 CATTAGGATAAAAAAGTTCAAAATTTGTAAAGTGAATTTGAATTTACAGAAATC
 25 ATTTATGTTAAATACGTGAAGAGCGCTTAATCAACTAAAAACGCCAAATCC
 TATTGTGTTTCAGTGGGATTTGGCGTTTTTATATGAAACATAATTATATG
 TATAGTTAATCATTCAAAAAGATTAATGTACGTTTTATATTGAATATTCAT
 ATTTCAAATGCATATGAATGACGTTTCATTAATATATACAACATAATCTTTA
 TATAACGATTTATAGATTTTAAGGTTTAACCTTCTGTTTGTGTTTAAATAA
 30 AGTAATATCCGTAAAAATACTGCAAGTGCAAGAAAAATCATCGCTGAAAAGT
 AAAAGGTATTATTTAAATTATTAGTAAATTGAGTAATTAGACCTCCGACTA
 GTGGGCCATCATTTGAACCGAAGCCTTGAACACTGTTGAACACGCCCCATG
 TTTCTTCTCTGTTTCGTTAGGATTAATATGCCAGCCATAAAGGTATTCGAAG
 CCGGTAAGAGGATACCGTACATTAGCCCAATAAAAAAGTCCATGGCCCAAA
 35 CTATATATATATTTGTAATTGTAGATAGCCCGAATATAAGAATTGTATATA
 GTATAAGCCACTAAAAATAACTCCATACATAAACCTTTGCTATTATTGTT
 CGATGATTTTGTGATAAAAAATAACATAGAGAAAGCACAGCCTATGCCACCAA
 TAATGATTGCTACTGTATATTCAACGGTTGATACTTTCACAACCTGCGTTG
 CATATTTTGAAGAATAGGTACAAGTGCTGCTATAGCTGCTCCTTGTAACA
 40 AGATACCCGGAAATAGAAATAAGATGACGTTGTGTAACATCTACAATTTGCT
 TTAATTGTGCTTTCACAGGTTTAGTATTGTAATTTGTTAAGTTGATATTAA
 CAAAATAGTATAGTACCAGGCAATAAGCACACCAAGGCCATTAAAAATG
 CAAAACGAGTAGGATGGAATTTAATAAGCAAGTTCATGATAACCATACCCA
 CCAATAAACCTAGCAACCATGAAAAGTAGACATAACCCATTTGTTTGCCGC
 45 GATTTCTTTTCATCTACACTAGATAAATGATAACCCAAATCGGACTCACAG
 CAATTCCTAACATAATAGCACTGAAAATAATTATGAATGGTGATGCCGGGA
 ACCATATCACTAAAAACAGCTTATAAAGCAAGTAAG

50 Sequence 3372

step.1002b09.cons.ok

TTGCTAATAAATATGGCAAGGTTAAGGCGATTGCGAGTGAAATGGTTCCAT
 GTACACCACATAGAGTCATGATTAGTGATATAAGCTACGCTTAGGTTTAC
 TCTCAGTAACTTTATCTTCATCATTTTTAGAAATCATTTTTTGGAATGGAC
 55 TGACAGATAAATAGAAATAAGGATAAAGTACATATACCCACACAAATCTAA
 AGAGATACACCGCAAGAGCGACTAGCAATGTGATGACGATTAGAATAATA
 AATTATGCGGTTTCAGTTTAAATAATTTTCACTATTACCTCAGGTACTAAAA
 AACCTAAGATTGAAAACACAAAACCATTTAAAACGTAGCCTAAAAATCTCC
 ATGTGTGATTATAGCTCATCTGTAATTGTGTTCTCGTTTGAGCTATACGGT

CACGTTCAAATCCATGACGAGTCCCGCAACTACTGCTGCAATAATCCCTG
ATGCATGGAAATAATTCAGCAATCAAATAAGTGACAAACGGCGTCAGCAATT
GTATAAATGTAAACATGTTAATGTTCTCAATGCCACGTCTCATTAATGTTA
AACGAAATCTCACTAAAGCCATGCCTATAATTAGACCAACAATAGCACCTC
5 CAATAGAGGCAATCAAAAATTGTTGAATCGCATCAAAAATGGAAAAAGTTC
CAGTAATCAATACACCTACGGCAATTTTAAATGAAATGATACCTGCAGCAT
CATTAAGTAAAGATTACCTTCAAGAATGGTCATTGAACCTTTTGGTAAAA
CCTTACCTTTTCGTGATTGCTTGTACCGCAACAGCGTCTGTAGGACAAAGTA
TAGCCGCAATTGCGAAAGCAGCACCTATAGGTAATTCAGGCCAAATCCAAT
10 GAATAAACAAATCCTACGCCATATAACGGTAGTTATGACAAGACCTAGAGCCA
TCATCATAACAGGTTTGTATATACTTCCTTAAATGTACTCTTGAAACATTAA
CACCCTCAACGAATAGTAGTGGTGCAATCAATGTAACCATAAATAATTCAG
AATCAAAATTTGAATTCAACTGGAATGGGAGTGAGGTATAAAATCATCCCAA
GGAAAAATTTGGATAAAATGCAAGAGGTACTTTAGGTATGAAAGTATGTACAA
15 AGGAGCTCACAATAACTAGTGCACAAATATGAGTAGAGTTTCAAATATCT
CCACACTTTACCTCTTTTATTAAATTTTAAATTGCAAAATTTAGTCATCAAAA
CAATAAAAAGTTGCAAGGTTATCTACTGCAACACTGAAGAAGTGAATAAT
AACAATAGAGTCATTCTGTAAAGACATAACTATTAACACTCAATCATGCC
ACAAATATACTATTCAAACTCAATTATTTTCTTTATCAGTGAAATCAAAC
20 ATGAACATTTAAATTAACCTTATGAAGCTCGCTTCTTCGAACGAGAAAAATGC
TTATGTGTAATTTTCAATGAAAAAATTAGTGACAACCTGTCATTCAATATAT
TCTTGTTTACAATTTTATCATTATTAACCAGATTCTTAAATAAATAGCTT
TTATATCATTAATATCTTCCACCTCAGATATACAAAAAAGTACTAGCATATA
AAAGATATATCTTTTATATGCTAGTAGTTTAAAGCTTAAACAATCAATCG
25 TCCATTTTATTAATATATCTTTTCTTTTATAGTATAATTTTCTCGTCTT
TCTTCTCGTAGTTTTCGACGTTTCTTAAATCTTCTGGCGTATTAACTCT
TTTTGGCTTAATAAGATAGTAGATTGCTTAGATGCATCGTCTTTCTGTAC
ATATATGCACCATTACGGTAGGCCGCTCTACTTATCAAATGCATTCACAACC
GGAGATGTAAGATTGATAAAAACTAGTGATAATAATAATCTGACACTGAAA
30 AAACCTGAATTACAATAAAATAGATCAGTACCAACTACAGTTAGTAAT
ACTGACAATGTAGAATTTTTCGTTGAGGCGTGACTTCTTAGAAAGACATCT
TGAAATTTACAATCCCTATTGCACTAATTAATGCAATAATACTTCCTAAG
AAAATAAGTATCTTATTATCTAATTAAAAAACAATAGAATACTTAGTTTGA
CTGTTCAATGATTAATAATATCTTTTAAATCTTCCCACTAAATACGAATAAA
35 AGCATATTTTACTGTTCTTACGCATTCTTCCGATTAGCTTTTCTGTATCG
TTTTTTCATPCAATAATAATTTTCTTAACATCTTCTAAGTAATCATCTTCAA
TCATTGGCTTTAAAGTATTTCCATCAATAGCATTTAACATTACATCTTCTT
TTTCCCTCAGTCTGTTATAAACTGCTTTATCGATAAAACCTTTATGGGCTC
TATCCCTTCAGACATCAAATAATAATACCTTGTATATTGATTATTATTTA
40 ACCCTAAACGATGTATTCTATCACGTGATTGCAACATAAACGTTAAATTA
AGTTATATTCAAAATATATTGCATCATGTACTGTCTGATGTAAAGATATGG
ACTCGCTAATGTATTAGGATTAGATATTAGAAGTTGTGCATTTCATTCC
TAAAATTATTGATCATATCTACCCTATCTTCTTTAGGTGTTTCTCCATAAA
TCAATATTGAATTAATATCACTTTCTAGTAACCTCTTATTGATTTTATTCA
45 TTGTACCTACAAACAATCCCCATACTAGAACTTTTTTTCTTGGCTAACTA
AGTTTTCAATCAATTCAATACCTTTTTTCGAATTTTGGAGAAGTTACATTTT
TTAAATCAAAAGAATTATAAATTTGTGTTTACTTGATTTTCTTCTTCTT
TATTTAATGCTTTATCCAAATCGAAATTTAATTCATCATTTCAACATTCTTA
GTTTCGCTATAAATTAATTTCTTTTGCAATAATCTGGGTTAGTTGAAGCTT
50 GTAGTAACCTTATATAAATCGCTAGTATCCAGATTCAATTTTTCGTATATCG
CTTTTGCTAATTTCAATTTGAATATTACTAGGCTTAACACATAATATAATAT
CATTTTCTGCTTGAGGCACTTCCAAATCGTCTTATTTGTACGCCAGAAAA
AAGGGTTCAACTTGTCAATTAATTTTCAATTAGGATCAGGATTTTGTAAATCAG
CAACATTTCAACCAAAAAAATAGTATCATACTCATCTTTATATAAAATATTTA
55 AGAAGTTAAATATGTCTTGATAACTATTTGGAATTTGGGGTACCAGTTAAAA
CGTATCTGTAATAAATTTTAGGACCTAAAGTTAATGCATAACTTGCTCTGC
TACTATTAATACCTTTTATCCCTATGAACCTCATCATAAACTATCATTGTCT
TATCATTAATAAGTTTCAATTTAAAAACCCCTACATACTTTTGTATTGCTTCAA
AATTCAAAACAATGACATTTGCACTTCCCCAATCTGTTCGTACTTTACCTA

AATCATTATATTTTTATCTCTTAGGTTTCATAAAGTGAATTCCTTTTAT
CTTGAAAACTTCAATAAATTCGAAACGCCAAGCTTCAAATGCATTAAATG
GAGAAATACTATTAATTTATCAACTTCCGCTTTATTTCACTA

5

Sequence 3373

step.1002b11.cons.ok

GCCTTCTTATAAAACGATTAATCGATTCAAAAACAAATAAAAAAATGAA
GAATTATAGTTGGGAATATTTCAATGTCCAAATTAATCAAAGCTTTCAGA
10 ACGAAAAGCAAAAACCATCTACAGTCAAAGAAAAATGATGTGGAGTCTGT
TTTTGGAATTATGAAGCCTATTTGAGTTTCACTCGAAAGTCCGTTTCGAGG
GATAAATAAAGACAAAAGAGAACTAGGACTTGTGCTAATGACACTTAATAT
AAGAAAAGTACCAGCTCAACGAGCTGAAAATTATAAAAAATCATAAAAAAG
ACAATTTTATATTATTTCAATAGAAAATTGCTTTTATTTACTTATCCTGGA
15 ACTTTATGTTCCGGACTCTTTCCTACGATTAGTCAAGTACTTATAATTAAA
ATAAGCATTACTTAACATAACCGTTCATTTTAATTTTTTTGAATAAGGTGT
TAAACTATATTTGATTGTACGTTTGGCTAATGTGGTGACATTTGATAATCG
TACAATTTATAAATTCATTACAAGATAATGAATTGTTTTTAATAATCGATTT
ATACAAGCAATGATGGCAGTCTTATGAGGTTTCTCATTAGGCTGCTTCTT
20 AGTTTGTAGTAATAATCGACGACATGATTGTCATAATGATGCTGCCCTCTT
ATTATATTCATAATCACCCAAAAATAAAGTTTTCCTCGCTTTTTTATTACCA
CGCTTGTGATGGTATCTCTACAGTGTGTATGACCTGATTGATATCGTTTG
ATATCAATGCCAACAAAAGCATTGAGTTGTTTTATTGATTAAATCGCTTA
ATATCACCAATCTTCCCAATAATCATAGCTGTGCTTAGCTTACCAATACCA
25 GGTATCGAATGAATATTTTCAAATAATCGAGTTGTTGTGCTAATTGAATC
ATGGCATCATCTAATTGTTTGATATGATGAATAGATTGTTTTAATTGTTGA
ATAAGTAAGCGTAATTTTTTCGACTAGAAAGGAATGTCTATCGACATTAGGA
TAGCTTTCTTGAGCAATCACCCCTAATTGAAGTGCATATTTGTAGCTTTA
TCCATTGACATTTCCCTTATCTGTAGAATTGAATATATGTGTAATCAGTACC
30 TCCTTGTGATATCAAGAACCATTGTCTGGATGAGTAAAGATTCTCGCATG
TTGAGTGAATGATTGAATATCGACTACTAAATAATCTTCTAAACCAGGG
AATGTTTGATGGAGTAATTCAGGATCTGAAATTTAAGTCGATTTGTGTTCA
TTCTCGATTTCTAGATGAAAACGACGCGTTCTCTTAATTCAAAGAATATT
AACTCATGTATAGGTAAGTTGTCTGTTTGTGTTAAGCGTCGGTCCTAACAA
35 GCAAGCTTATGAGCATCTGCCGTATCAGTTTTCCATGATCTTAGAGCGCTC
GTTTTAAATTTGGCTTCTAACGGATTCAATTTGAATATAGTTAATTTGATTT
ACACAACAAAATCGTTCCATACCTCTTGAATAGATACCTGTAGATTCAAAA
ATGAGTTGTGGGTGGTCTAAGTCATTCAAATACTTGAGTAAATAGTTGTAA
CCATTTTTATTATTTTGGATGAAAACTCTTTTTTGAATTTTTTCATTTTTA
40 TAATGTGCAACTACACTACTTCTTTTACTAATATCAACACCTAAGTAATCG
ATAAAAAAACCTCCTTTGAATAATTGAGAAGCTAAAACTTTACTTAACCT
TTCTCATTTCATTTTCTTATACACGGTTTCAAGAACCCAACATACTACAAA
CGAATTTCAAAGGCGAGAGTAAAGCTGACTTGTTTTTTATACGGATTAA
AATCCAAGAGTCTGGACAGTCTACTTCTCTCTATAACTATAAAAAATAGCT
45 ATGAAAAAATCTATCGTCATAGATTTCTTCATAGCTAATCTTAGTATGTTT
TTATTTTATTGAAATTAATTGTTTATTAGCGATATAACTTGTCAATGATAG
AGTCGTTATGTCAATTATAAATGTAAGGGCTTACATATATGTTTTGCATCAA
CATTTAGTCTGTTTCATACAAATCAATTTTTTCGATTATGACGATATTGATT
CATTTTATTGTTATATTAATAAATAAGTTTGAGAGTCGATAGAAATCAGTT
50 ATGATATTGTTAATAAGTTTTTATTCGAAAACGAAGGGGAGATTTTATGAC
TCAATCTGAAAAAATTATTAACCTTAACAAATCACTATGGGGCGCATAACTA
CGTTCCACTTCCAATTGTTATTTCTGAAGCAGAGGGTGATGGGTGAAAGA
TCCGGAAGGTAACACGTATATGGATATGCTTTCCGCCCTACTCGGCAGTGAA
TCAAGGTATCGACACCCAAGAAATTATTCAAGCATTGAAAGATCAAGCAGA
55 TAAAGTCACTTTAGTATCACGTGCTTTTCATAGTGATAATTTGGGTCAATG
GTATGAGAAAAATATGTAACTCGCAGGTAAAGACAAAGCATTGCCATGAA
TACGGGAGCAGAGGCGGTTGAAACAGCTTTAAAGCTGCTCGTCGTTGGGC
TTATGATGTTAAGGGTATTGAGCCGAACAAAGCTGAAATTATCGCTTTTAA
CATAAAAGCTATCAATAGCAATACAAATACATTGACGACCAAATTTCTCGC

TTGCTTGTCGAATTAATTCTGGATGTTTTAAAGCGCTTGAATTGAGTGATA
 TTTTATCTGCTCCGTGATTTAATAGTTGCGTAATATCATCTAAATTTTGAA
 TCCCTCCTCCTACTGTTAAAGGGATAAAATAATTGTTTTGCCGTGCTTCTA
 TCACTTCTATCATAAGATCATGTCCTGCTTCCGTCTCGAAATATCAAGAA
 5 AGACTAGTTCATCTGCACCGGCTCATTATAATAAGAGCCAAATCAACTG
 GATTACCGATATCTCTTAATGACTGGAAGTGGATACCCCTTACGACGCGTC
 CATCTTTAACATCTAAACATGGAATCACTCTTTTTTTAATCAAGATAATCC
 CTCCCAGAATTCAATCCAGATGTGCTGCTTTTCCCTACAATAGCAGCATGAAC
 ATTTAACGATTCTAATCGAAACAAGTCCTCTTGATGTCTAATACCTCCTGA
 10 AGCAATTACAGGCAACGATGTATAAAGTGCAGAGACGACCTGTCTTTAGAG
 CACAGTGGCGATGATATC

Sequence 3374

step.1002b12.cons.ok
 TATTTAGCTGCTATCAAAGAACAAAATGGTGCAGCAATGAGTTTAGGACGT
 ACTTCAACATTTCTTAGATATTTATGCTGAACGTGATTTACAAAATGGTGAC
 ATCACTGAACAAGAAGTTCAAGAAATCATTGACCACTTCATTATGAAATTG
 CGTATCGTTAAATTCGCGCGTACGCCTGAATATAATGAATTATTTCTTGGA
 20 GATCCAACCTTGGGTAACCTGAATCTATCGGTGGTGTAGGTATTGACGGACGT
 CCAATGGTAACTAAAACTCATTCCGTTTCTTACACTCATTAGATAATTTA
 GGTCCAGCACCAGAACCAAACCTAACAGTGTTATGGTCTACTCGCTTACCT
 GAAAACCTTCAAATCTATTGTGCTAAAAATGAGTATTAAAAACGAGCTCAATC
 CAATATGAAAATGATGACTTAATGCGTGAAAGCTATGGCGATGATTATGGT
 25 ATCGCTTGCTGTGTATCTGCCATGAAGATTGGTAAACAAATGCAATCTTC
 GGTGCACGTGCTAACTTAGCTAAAGCATTACTTTACGCTATCAATGGTGGT
 AAAGATGAAAAATCTGGTAAACAAGTTGGGCCAAGTTATGAAGGTATTAAA
 TCAGACGTACTAGATTATGATGAAGTCTTCGAAAGATATGAAAAATGATG
 GACTGGTTAGCTGGCGTATATATCAACTCATTAATATCATTCACTATATG
 30 CATGATAAATATAGCTATGAACGCTTTGAAATGGCTTTACATGATACAGAA
 ATTATTCGCACAATGGCAACTGGTATTGCCGGATTGTCTGTAGCAGCTGAC
 TCTTTATCAGCGATTAAATATGCACAAGTTAAACCTATCCGTAACGAAGAA
 GGTCTTGTAACTGACTTTAAAAATCGAAGGCGACTTCCCTAAATATGGTAAT
 AATGACAGTCGTGTTGATGAAATTCAGTAGATTTAGTTGAACGTTTCATG
 35 ACTAAATTACGTAGCCATAAAACATACCGTAATTCTGAACACACAATGAGT
 GTATTAAACAATTACTTCAAACGTTGTTTATGGTAAGAAAATGGTAACACA
 CCAGATGGACGTAAAGCTGGCGAACCATTTCACCTGGCGCAAACCCCAATG
 CATGGTCGTGACCAAAAAGGTGCATTATCTTCACTAAGTTCAGTAGCTAAA
 ATACCTTACGATTGCTGTAAAGATGGTATCTCAAATACATTTAGTATCGTA
 40 CCGAAATCACTAGGTAAAGAAGAAGCAGATCAAAATAAAAACTTAACAGT
 ATGTTAGATGGTTATGCAATGCAACATGGTCATCACCTCAACATTAACGTA
 TTTAATAGAGAAACATTAATTGATGCAATGGAACATCCAGAAGAGTATCCA
 CAATTAACGATTTCGTGTATCTGGATACGCTGTAACTTCATTAAATTAACA
 CGTGAACAACAATTAGATGTTATTTACGTACATTCCACGAATCTATGTAA
 45 TAAATTTAAGGTGGGAGCAAAATGCTTAAAGGACACTTACACTCCGTTGAA
 AGTATGGGCACGTGTCGACGGGCCAGGACTTAGATATATATTATTTACTCAA
 GGTTGTTTGTAAAGATGTTTATATTGTCTAATCCAGACACTTGGGAAGATT
 AACGAACCATCAAGAGAAGTGACGGTTGATGAAATGGTAAATGAAATCTTA
 CCGTACAAAACCTTACTTTGAAGCTTCAGGTGGTGGGTAACAGTCAGTGGT
 50 GCGCAACCATTACTACAAATGCCTTTCTTGGAGCAATTATTCAAAGAATTA
 AAAGCGAATGGTGTTCACACATGCATTGATACTTCTGCGGGATGTGTGAAT
 GATACACCAGCATTTAATCGTCATTTTGATGAATTGCAAAAGCATACAGAT
 TTAATCTTATAGATATTAAACATATTGATAATGATAAGCACATCAAATTA
 ACAGGCAAACCTAACACACATATTTAAAGTTTGCACGTAAATTTATCTGAT
 55 ATGAAACAACCTGTTTGGATTAGACATGTTTTAGTACCTGGTATTTCCGAT
 GATAAAGAAGATTGTATAAACTAGGAGAATTTATTAATCTTTAGATAAC
 GTTGAAAGATTGTGAAATCTTACCATATCATCAACTCGGTGTGCATAAGTGG
 AAAAATTTAGGCATCCCTTATCAACTCGAAAATGTTGAACCATCTGACGAT
 GAAGCGGTAAAGAAGCTTATCGCTATGTTAACTTTAATGGCAAAATACCC

GTAACATTATAGAAAGATAAAATATAGTAATAGCCTGAGACATCAATCGTTG
 TCTCAGGCTTCTTTACGTTTAAATCAGTCGATAAACTGAGTATAAAATATAT
 CTTCAATTAACCTTTATTTAATAAAAAATATGATTTCATAAGTTTAAATAGTTTA
 TTTTCTAGATATGCTTTAAATACATCTTTATAAACATCGATGTAACCTAA
 5 ATACATATCTTTCTAATAAAATTCATCAACTTGATGCGCCATAATAGATTC
 ACCAGGTCCAAAGACAGCAAAATCCACGTTATTTTCATTTGTTCCTAGGAA
 ACTAGATGCATCAGTTGTACCAATCAATGCACTCACAACAATGTCTTCATG
 TACATAATTCGGTGCAATACGTGTAATATTTGAATAAGAGGATTATCACG
 ATCACTTGCCACTGGATCGTGACTGCTAGGTATATCTACAGTTAAATAATC
 10 TTCGCCACATGACGAATGACTTTTCAAATAAATCCTTCACGAAAGTACT
 GTCATATTTGGAACAGTTCTTACATTATATTTAGCTGTCGCTTTATGAGG
 AACAGAATTAACCTGTTTACCACCATTTGAATACAGAGTTTAAATTAACAA
 TCCAGAGTAGATATGTGATTCTTCTTCACCAATTTTCTGTGGAGATGTTT
 CTCAATCATTTGGAACAGCGCTTAACTCGTGTACTTTATCATGTTCTTTAAT
 15 ATTTTTATATTTCTGTTTCATTTTCAATTTACAAAATCACTAAAATATCAAC
 AGCATTTGTACCTAGGTGTCGTCATGGAACATATGTGCAGCCTTACCTTTTCGC
 TGTGACTACACAGCTCATAGAACCCTTTATGAGCATAGTAAGCGATATTACT
 AGTTGGTTCGCCAATAATTAACCACCTCACATCATCTAAGTAACCTTCATC
 AGCTAGTAATTTGTGCACCATATTTGTCCGCTCTCTTCTCCTGTAGTAGCTAA
 20 TAAACGTATCTGTACCTTGCTTTAATGCGTTTGATTGTTTAAATTCATCAT
 CGCGATGACCATCGCCATAAGACCACCTTTTCATGTCTAGTGGTACCTCTGCC
 GAATAACTTGCCATCTTTATCTGTGAGTTCAAATGGAGGAAAAGTCCAATC
 ATCATGGTCACCAGCATCTACAACATCCATGTGTCCACTGATTGCTAAGAC
 TGGAGCGCCACTACCAATTTTCACTAACTAAATTAGCGCGTGAATCATTAAC
 25 TTTAACAATCTTAGAATCAATATCATATTGACTTAACAGGTCTTTTAAAGTA
 CTCACATACTTTAATTTTCATGATCATTTCTCAGTTTGAATCTTGACGATATC
 AGCTAATAATCTTATTTTGTCTTGCTCACTCAATACAGTCATTTTCACAGTC
 CTCCTCATTTTGTAGTAATGATTTTCACTTTTAAATAAAAAATAAAAAATCCT
 TGTACACTTAATATTATCAAAAACTTTTATTAATTAAAAATTATATGCTAA
 30 AATAAATGGATTTTAAATAAAAAATAAATATTAACCTAACATTTCTTTATAA
 AAAACAACAATGTATAGAAATATTAATAAATTTTTCATTTATTTATAGTTTGA
 ATCGTGTATAAAATGAAAAGCTATGA

35 Sequence 3375
 step.1002c01.cons.ok
 TCAACATTATAATATTGATTAATTAATAACAATTGCAAGTATAATTTAAA
 TATTTCTCTAATATACAGTGTCAATTTATTTTATTCACATAAGAAAATAGC
 TATGAAGAAATCTATCAATTTAAATTTCTTCATAGCTAATTTTTTTCATTT
 40 AAATTTATTGACGGCTTGAAAAATGAGTCAAAATCATCAATAACATCAAAT
 TGCAAAATATATTCCTTTGGTAATGGATTGACCATTAAACTTAAATTCGAATT
 CTATTCCTTTCTATACAATGAAACGGGTGTCATACATCATCGGTAACATAAT
 TATGATAGATAGAACTTGTGGTTCTTTATCGTCTTTAGTTTACTAATGA
 GAGCACGTGGAGTATTTCCATCTTTGATTCTAATTTTCATACTCATCTAGTT
 45 TATCAAAATATTTTTTCGGCTTGCTCTGTAACATATTGTGTAATACCTATCG
 TTTCTGCCGTGTCGGTAATAATCTATAGGCAAATCAACTGTAAGTTGTTTAG
 CTTTTTTATTTACGAATTTAACCCTTACCAACTGCTTGTTGTGAAGTTTGAAA
 AATACGATTGCAAAATATCATTTAAACTGTTTAAAGTTATTATTCAACGTTT
 CATCATAATCAGCTGCAGTTGACGAAGGAATTAAGGCTGCTTTTTTCATTAA
 50 TATTATCCCATGAGTTAATTTTAGTTTATACCTCTTCAACCGTAGTACCAA
 CTATAAATTCACCTGGTGAATGGAATCTTGACTTGATTGTTTATAGATAG
 CAAAATGAATAGGAATATCTTTCAAATCACTATTCTCACGTAAACGAGAAA
 GCATTTCACTAGCCATCTGTTTACCTTGCTTTTCAATCTCTTTATCAGATA
 AATCTTTACTAAATGTTTTCGCCATCTTTCTCTTTTTTGTAAATAATAAACAC
 55 TATTCATAGCTAAACCAATTGTCATCCCTTTTATATTTTACCTTTAGAAT
 CACTATTTCCATAAAAAATCCTGCTCGAGTATATTTGAAAGATAGGCTGGAG
 AATTTTTAGCTATTTTCTCTTCATCTGTTTCAACCATTTGTGAGATGGATTGA
 GTCCAAGATTCTCATTAGCATTTTGTCTTTTCTTTTCTTTTTCGCTCATCT
 TGTCAATTTCTTTTTTCGTATACCTCGGATCTAAGTATGCATTAATCGTTT

TTTTATCTAAATATTGTCCATCTTGATATAAATACTTATTTGTTGGAAAGA
 TTTCTTTACTTTAATTTCTAGTAAACCACTTTCAAATCTTCTCCATTATAAC
 CATTGTCATATTATCTTGTAATAATCCACGAGCCTGGCTTTCTTTGAAGG
 5 GTAATATAGTCCATATAGTTATCACCTTGAACTTTTTATCAGTCGCTATTT
 GTTTCATTGATTTTTATTATGGTTATCCTTATGTTCACTTTGTTCTTTAT
 CTGATGAAGTTTGTATGTCCATCTCCGCAAGCCGTTAATAATAACAGTA
 TCGACATGAGTAAAAATATTGTTGCTTCATTACGTAAGTCTCTAATTAT
 TAGATTCCATTTGTTTTTCAATAAATGCTGCTTCAGTCCAAATTTAGTA
 10 CCATACCTCTCAGCTTTGGCTAATTTAGACCCGTCATCTGCTCCAGCTATG
 ACAATATCAGTACTTTTAGTCACGCTGCTTGTAACCTTTAGCACCTTGCATT
 TTCAACCATTCAGATGCTTTCATTTCTCGTCATTTGCTCGAGTTTCCCTGTT
 AATACAAATGTTTTTCCCACTAAAATCAGGATGACCTTCGATTTTCAGTTGTT
 TTAATTCCTTTATAAGACATATTAACATTTTTATTACTTAATTTTTCAATT
 AATGAACGAATATCACTATTTTCGAGATATGTTACAACAGATTGTGCAAGT
 15 TTATCTCCAATATCTTGAATTTCAATTAATTCATTTTCAGTTACTTTAAAA
 AGTTGATCCATCGTTTCATATCGCTCAGCAAGTACTTGACTAGCTTTTACA
 CCTAAATGCTAATACCAAGTCCAAATAATAAATGCTCTAATGACTGTTCT
 TTAGATTTTTCTATCGCTAATAAAAGATTATCAACTTTCTTTTCCCATTT
 CGCTCTAATGGTAAATAAATCTTCTTTTCAAATAGAAAATATCTGCGACA
 20 TCTTTGATTAAGTATTTTCGTATAGCTGATGAATAATTTAGTACCTAAA
 CCATCTATATTCATCGCTTGTCTTGAAACGAAATGTATAAGTCCCTCAATA
 AGCTGTGCTGACATTTTGGATTAATACAACGTAAAGCAACTTCTCCTTCA
 ATACGAACTAATTCATGTCCACAACCTAGGACAATGTGTTGGCATATGATAA
 ATTTCCGATTCGTTAGGTCGCTATCTAAAATACCTTTTACAACCTTCAGGG
 25 ATGATGTCCCCGGCTTTTTTAATAACAACACTATCTCCGATACGTATATCT
 CTTTCATGTATTAATCTTCATTTATGAAGTGAGGCTCTTGAACTGTAGTA
 CCAGCTACTTTTACAGGTTCTAGAATTGCAGTTGGTGTCAACAACCCCGTA
 CGCCCAATACTTAGCTCAATATCCAATAATTTTGAATAAATCTTTCAGCT
 GGAAATTTATAAGCAATCGCCCATCTTGGAGATTTTGGCGTATAACCCATT
 30 TCCTCTTGTGAGATAAATCGTTAACTTTTATAACAATACCATCAATATCG
 TAAGATAAAGATCCTCTTTTGTCTTGCTCATTCTCTATATAATTAAGTACG
 CCCTCAATATCTGATACCTGCTTCACGTTCTTGGTTAGTTTAAAACCTAAT
 TGGTCCAATTCCTCTAGCGCTTCACTTTGTGTTGTTGCATTAAACTCGGTT
 AGGTCAATTCACACTATATAAGAAGACGCTTAACCTTCTTTTCGCAGCTAGT
 35 TTAGAGTCAAGTTGTCTTAAAGAGCCTGCAGCAGCGTTTCGTGGATTGCA
 AAAGGTTGTTCAACATTTTGTCTTTTTCATTATTCAAATGAATGAATGAA
 CGACGTGGCATATAAGCTTCCCCACGGACTGAAGGATATCAGGTTACGACA
 ACAACATGTGGTAAAGAAGCACTAAAATTATTATCTTCAGATATAGATATT
 40 ATGATTTTAGACATTATGATGCCGGAAGTTAGCGG

Sequence 3376

step.1002c03.cons.ok

GGTGTGTAAGAGATGTAAGGAATAGACGTTAGTGGCCCATTTCCACGTAT
 45 GACATATGCAGAGGCTATGGACCGTTTTGGTTCAGATAAACCTGACACTCG
 TTTCCGTATGGAACTTATCAATGTGTACAGCTTGGTAAAGAAATGAATTT
 TAAAGTTTTAAAGATACGGTAGATAACAACGGCGAAATTAAGCAATTGT
 CGCAAAAGACGCTGCAAAATAAATATACACGTAAAGACATGGATGCATTAAC
 AGAGTTTGTAAATATATATGGTGCAAAAGGATTAGCTTGGGTAAAGTTGT
 50 TGATGATGGTTTAAAGTGGCCCAATTGCTAGATTTTTCGAAGATGTTAATGT
 TGAAACACTTAAACAGTTAACAGAAGCTAAACCTGGAGATTTAGTAATGTT
 TGTAGCTGATAAACCTAATGTTGTGCTCAAAGTTTAGGGGCTTTAAGAAT
 TAAATTAGCAAAAGAATTAGGTTTAAATTGATGAATCAAAATTAATTTCTT
 ATGGGTAAGTATGGCCGTTATTAGAGTATGATGAAGATGCAAAACGTTA
 55 TGTAGCAGCACATCATCCATTTACTTCACCTAAAAGAGAAGATATCGAAAA
 GCTAGACACTGAACCTGAAAAATGTACAAGCCAACGCTTATGATATGTTCT
 AAATGGTTATGAACCTGGTGGTGGTTCTATAAGAATACACGATGGTGAATT
 GCAACAAAAAATGTTTGAAGTATTAGGATTTACTAATGAACAAGCTCAAGA
 ACAATTTGGTTTCTTATTAGATGCTTTTAAATACGGTGCTCCACCTCATGG

TGGCATCGCGTTAGGTTTAGATAGACTTGTGATGTTATTAACAAATAGAAC
 AAACCTTGAGAGATACAATTGCATTCCCTAAAACAGCATCAGCTACATGTCT
 TTTAACTGACGCTCCAGGAGAAGTATCTGATAAAACAACCTCCAAGAACTCTC
 ACTAAGAAATCAGACACTAGGGCCCCAACAAAGAGAATTTACAAAGAAAAT
 5 CAACAGGCAGAGCAAGTTGGGACTAGGGCCTTAATTTAATGTTTCTTAAGA
 AATTAGCACAGAGAGCTTAATTGTACGAATTGAATATAAAACATCGGTGTG
 TAATATAAATTAATCAGACATTATATATAATTTTGGTAACTAGTTTAGAAT
 ACACGAGAAAATAATACACATTTGAGTTAAATGATAAGTCTTTAAAAATTTA
 CTTCTCTTTTAAATTTGTGGTATGATATTTTCATAAGATAGTCATCTGTGG
 10 TGTTCGTAAGTTTGTCTTTTATTTGGGCCTAACACTCTTTGATCAGGGAGC
 CCAATAGGTTTCTCGCAGCGCACACGCCTCTATAGGAGGACTTGCAAAAC
 GAGAAACAGGGCACCCACCTGTATATAGCAGGCCGAATGATCAAGCTATTT
 ATAACACGGCATCAACGGACTCTATCGGTACGCAAGACTTTTGTCTTGCG
 TATTTTTATGTATATAAATTATAAAGAAGCATAGATATAATGAATTGGAATC
 15 AAAGGAGCTAAAGATGAAACATCAATTTTCAAGGAATGAATTAGCAATAGG
 ACAAGAAGGGCTGAACCTTACTAAAAATAAGACTGTTGCAGTTTGTAGGTGT
 TGGTGGCGTCGGGTCATTTGCAGCTGAGGCATTGGCTCGGACTAATATAGG
 GCACATCATACTTATAGATAAAGATGATGTCGATATTACAAATGTGAACAG
 GCAAATTCATGCACTGACTTCAACTATTGGTCAAAGTAAAGTCACGCTAAT
 20 GGAAGAAAGAATCAAATTAATAAATCCCGATTGTAAAGTAACTTCTTTGCA
 TATGTTTTATACCGAGGAAACATACAAAGATATCTTCAATAATTATGATAT
 TGATTATTTTATTGATGCAAGCGATACAATCATTATATAAAGTTTCATCTCAT
 GAAAGAGTGTTTAGAAAGAGGAATTGAGTTAATTTCAAGTATGGGTGCAGC
 AAATAAGACTGACCCGACACGTTTGGAAATTGCAGATATTTCAAAAACACA
 25 TACTGATCCTATGGCTAAAGTAATTAGAAATCGTTTAAACCGCCTTGGTAT
 TCGTAAAGGTGTTAAAGTAGTATTTCTGATGAAAGTCTTATTGTTATTCTG
 CGAGGACGTAAAAGAAACAGTAGGAGATAAAAAATGCAATCAATAGAAAAGG
 GCAAATGCCTCCATCTTCTAATGCATTTGTTCCAAGTGTAGTAGGCCCTTAT
 TTGTGCAAGCTACGTTGTCAACGATATTTTAAAGATATACCTGTAAGGCG
 30 AATTAAAGATAAAGGACAAAAATTAAGTTATCTATCTCATTGCAAGAAAAT
 AATCATCTTCTATTTTGTGAAATTTGATTCTAGTTTAAAAAAAACAAAA
 TGTATTTTAAATAAATCTACAACAACCTCATAGTAAAAAACCTGAGGCATC
 AGTATTGTCCAGGTTTTTTTGCATCTAAGATATTAATTTCAACTACATATT
 TTCAATGAATAAAATATGTTTTATTTCGTAAGTTATCGTAAATGTTTTTA
 35 AATTGTTGTTCACTTTTGTGATGTGGTTTTTGGTTTCATAATACACTCTGTTT
 TTTAATTTTCTGGTAAGTACTGTTGTACAACATGACCATTTTCATAGTTA
 TGAGGATATTTGTAAACCTATAGCCCGTCCTAACTTTTTCAGCACCTGAATAA
 TGACCGTCTTTCAGATAATCCGGTATTTGACCTATATGTCCATTTCTTATA
 TCGCCAAGAGCTTTATCTATTGCTGTGATACCAGAATTTGATTTAGGTGAT
 40 AAGCATAAATCTATTACAGCTTGACTAAGGGGAATTCGTGCTTCGGAAAA
 CCTAGACGTTTCAGCTGATTGTATTGCTGCTAAAGTCCCTCTGACCAGCATTA
 GGTGATGCTAACCCAACATCTTCATAACTAATTACAAGTAATCGACGTACG
 ATTGTAGGTAAGTCACCAGCTTCAATTAGACGTGCTAAGTAGTGGAGTGTCT
 GCATTGACATCACTACCTCTAATTGATTTTTTGAAGCGCTCATGACGTCA
 45 TAATGCATGTCTCCATCTTTATCACTTACAAAAGCACCTTTTGTGAAGCAA
 TCTTTTGGCGTCATCTAATGTAATATGTCTTTTCGTTTTCTTCACCAATATGA
 GCACCTTAGTACAGCTAATTCCAAAGCGTTTAAATGCACTCCTAACGTCGCCT
 TGGCTTTGCGTTGAAAAATATTCGATAGCATCTTCATCTACAAATAGGATGA
 TAAGTACTTAAACCTCTTTCCTTATCATTTATAGCTCTATCTAATGCCAAT
 50 CTTATATCATCTTGATCTAATGGGTAAAGCTCAAATATTTGTGCTCTAGAT
 CGTATCGCTGGATTTATCGCATGATAAGGATTTGAAGTTGTTGCACCTATT
 AATACAATTTTGCCATTTTCAAGATGTGGTAATAAAAAATCTTGTTTAGCT
 TTGTCTAAACGA

55

Sequence 3377

step.1002c04.cons.ok

ACATTATTGCGTTTTAATTGCAGGTATTGAGAAATGCAGACGAAGGACGTATT
 CAATACCTCAATCAATATTTGTCAAGACGTCGAATACGTCATATTGTAGGC

TATGTCCCTCAAGACATCGCACTATTTCGAGCATATGACTGTCATGGAGAAC
 ATTGAGTTTTCAGTCACTTTGTGAAAATCCTATTTTCAGATGAAACACTT
 CATTCTTATTTATCACAATTAAATTTTACTGATACAAAAGTGAAGTATCT
 AACCTTTCTGGGGGAAATAAACGTAAAGTCAATATTATGATAGGTCTACTT
 5 AGTCGGCCTAAAATACTTATTCTAGATGAGCCAACAGAAGGCATTGATTTA
 GAATCAAGATATGATATTCACAACCTTATTACAACAAATGACCGATCAATGT
 TTAATCATCATGACGACACATCATTTAGACGAAGTTGAAGCACTAGCAGAT
 GATATTAAAGTTATAGGTCAAATCCTTTTTATCATGATATTTTAGAAAAT
 AAAGGTTGGTCTTTTAAAAAATATGCAAATGCCCTTAGCTGATAATACGAAA
 10 TCTTAAGCTGTACTTTCACCTCCATCACTTCATCAATATTTAACATCAAAA
 GGCTTGCCTTTTCGTCGTATATTCTCGATTGAAAAGGCAAGCCTTAATAA
 ATTTTATACACATTTATTTATCGTGTCTAACTTATCTGTACACTAATCCAC
 CATCAGTTAAATGGATTGACCTGTAATATAATCAGAATCATTTGAAGCTA
 AGAATGATACTAAGTTCGCAACATCTGATGGCTCTTGATATCTTCCAAGTT
 15 TGATTTCTGAAGAAAATGCTTCAAACGCGTCTCCTATTTCTAGACTATCAT
 CAAGTTTCACCATTTCTTCATCAATACGATCCCACATTTTCAGTTTGGCAA
 CTCCTGGGCAATAAGCGTTTACGGTAATGCCTTTGTCTGCTAATTCTTTAG
 CAGCTGTTTGAAGTGAAGGAACGTACAGAGTGTTTGTAGCTGAGTAAGTGC
 CGAGTACTTCATAAGATTTCATGTCTGCAATACTACATGCGTTGATAATTT
 20 TACCTTTACTTTTTTGTAAATAAATTGGTTAGCAGCTGCTTGAATTCAC
 ACAATGTACCAAATACGTTAATTAATAAATTAGATAATTCCTCTTCAC
 CAATTTCTAAAATTGGCGTCACTGCATCAACACCTGCGTTATTCACCATGA
 CATCTAATTGACCAAATTCGTAACCTGCGAAGTGAACCTTCTCTCTGTT
 CTTTCTTTTGTAGATACATCACTCTTGAATGCGACAGCTTGATAACCTTTT
 25 CTTTGAAATCTTTCTCAGTTTCTAATAGAAGTGCTTCATTAATATCTTGTA
 GTACTATAATPAAACCCATCGTTAGCAAGACGTTCTGCAATACCTTTACCTA
 ATCCACCTGCTGAACCTGTTATAATTGCTGTTTACTCATATCTATCTTCC
 TCCTTATTTATCATCTACCTTTAAACATGGTGGTAGTAATAAGATTGGGA
 AATGTCAAATGTTGTTAATATATTTGTCGAAAAGTTGTTGTTTATGTCAT
 30 TATTAAATACAACCTCAAATTACTAACACTGCTTTTTACTGCAACAATATA
 CCTTTTAGATATCTACATCTTTTGTAGGAATTAAACAATACGCCGCAATAA
 ATGCGATAAAACACTAAACCTGCACCCACGATAAATGCGATAGTAGCTGCTA
 AATATGCCGGGTGATTTACAGATAATGCCGTATATATTGTTGTCACAACAG
 CGATTCCAAATGCTGCACCTAGTGTAGAAGTCATTTTGATAATTCCTGATG
 35 CAGTACCTGCTTTTTTCAGCTGGAACATTAGATACAGCTGTAGATAGCGCGG
 GTGTAGCAAAAAATCCTAATCCTAGACCTATAAAGATAAAGCCAAATACATG
 CCACAATATAATAATCATATTAGGCAAAGAAGTGAATGCTAAAAGTATAA
 TACCAACAGATAAATTGATACCGGACCTAACATCAATGGCAATTGCGGTCTC
 TTTTTTGCATAAAACGTTTACCAACACGAATCATTAACAACTACACAGCA
 40 TATATGGAATTGTAATTAAACCGGCTTGTGCAGCTGATAGATGTTTATCGT
 CTTGAACATAAATATTAATAATGCTAATGAACCAATATCCATGTTTACCA
 TCAAGTTGGCTAATGTTGTTCCAATATAAACATTATTTGAAAAATAAATTA
 AATCTATAAAAGGTTTATCTTGACGTTTTTCAAATATATAAAATGCAATTA
 ATGTCACAATGAATATAGCAATTAATATCAATATAAGAGGATTTAACCATC
 45 CAATCTATCTCCTTGCGTAATCACTACGTTAATGCTTAACATCATGACAA
 CAAAAATAATGATTCCAACGATATCAAATTTTTTATTATATGCAGACTCAT
 CTTTAGATTCTGGTATACCTCTAAGCAATATGAGTGAATTAAGTTTAAATA
 ATATAATAAGAGAATTTAACAAGCGTTTTGAAGAAAAATATAGTCAAAGTT
 ATGAATATAAAAAATTAGATTTTTCTGCTACTACAACCATTTATGATTGAG
 50 AAATATCAGAGTTTAAAGATTGGGTAGATGCAAAATATTTAGGTACTAACG
 TTGAAAAATACATTCAAACCTGAAAAAGATTTTATATGAAAGACCACCAG
 TTAGATATGATAGTGTAAACCTGAGTTAGAGTTGTTAAAAAGAAATTACG
 ATAAAAATGTAACCTGTATTGGTAATTTGCCATAAAAAAGCGATACAAGTTC
 CTAAATATACTGGTGGCACTACTACGCTGATTTTGTCTATATGATAGAAA
 55 CTGATGAACAAGATGCAAAATACCTTATTGTTGAAACAAAAGCAGAAAACA
 TGAGACTAGGAGATAAAAGTATTGGTGAAATACAAAAAAATTTCTTTAACA
 CATTAGATAATTTGAATATTAATATCAATTAGCTACTAGCGCGCAAGATG
 TTTATAATGAAATTAAAAAATTAGATGATTCAAAGTGAGGGAATATGATGG
 GGAAATCAGAAAAAATTCATTACTTGA AAAAGTCCAAGATGGTTTAGTAG

ATAAAACCGGAGCAAAACCTAAGTTACCAACAACCTATTACCGATTTAAATC
 AAGAACTTTAGAGGTATATAGAATACCACTAAAGTTTCTGTATTATAATG
 ATAGAAAATGGAAGAATTGCTTCTGTAAATATCCAGAGTAAGCGACGATATAA
 AAGTTGCTTATGAATTTGAAGATAATAATTATAATAAAAGTATTGAAAAATA
 5 TGATATATGAGGCTAATACCTTCTGCTTTAAAAAACACTAAAAAATCTATTA
 AAGATAAA

Sequence 3378

10 step.1002c05.cons.ok
 AGATATCTCACCATTAGGTGCTGCAGCTTTAAGTGGAACCTACACACCCTAT
 AGATAGACATCTGACTCAAGAACTTCTAGGATTTGCAAATCTGTATGAAAA
 TAGTTTAGATGCTGTCAAGTATCGTACTATATCGTTGAAACATTGCATCA
 TATTTCACTCACTATGGTACATCTATCACGTTTTGCAGAGAAATCATTTT
 15 TTGGTCGACTGATGAGGCAAAATTTATCACTTTATCAGATGCGTTTTCTAC
 TGGCTCATCTATTATGCCACAAAAGAAAAACCTGATATGGCTGAACTAAT
 AAGAGGAAAAGTCGGACGTACTACAGGACACTTGATGAGTATGTTAGTAAC
 ACTTAAAGGCTTACCCTTAGCTTATAATAAAGACATGCAAGAAGATAAAGA
 AGGTTTATTTGATGCTGTACACACACTTAAAGGCTCTCTTCGAATCTTCGA
 20 AGGTATGGTTCATCTATGAAAGTTAATTCAAACCGTTTAAAGTCAAACAGT
 AAAAAATGATTTTCAAATGCAACAGAAATAGCAGACTATTTAGTCAGTAA
 AAGTGTACCTTTTAGAACCCTCATGAAATCGTTGGTAAAAATCGTATTA
 TTGTATTATCAAAAGGTATATACCTATTAGACGTACCTTTAAGCGAATATCA
 AGAACATCATGAGAATATTGAGGAAGATATATGATTATTTAACACCTGA
 25 AAATTGTCTCAAGCGTCGCCAAAAGCTATGGTTCAACTGGTCAAGAATCAGT
 AAAACATCAACTAAAAGTCGCAAAAGCATTATTTAAAGACAACGAATCAAA
 ATAGTTATTAATAATAATAGCCATCTAAAATTACAAAAATTTAATTTTGTA
 TTAAGTTATCAAAGTTAGTGTAAAGTGTAGGGGGTTTCCGCCCTTAGTG
 CTGCAGCTAACGCATTAAGCACTCCGCCTGGGGAGTACGACCGCAAGGTTG
 30 AAATCAAAGGAATTGACGGGACCCGACACAAGCGGTGGAGCATGTGGTTT
 AATTGCAAGCAACGCGAAGAACCCTTACCAAATCTTGACATCCTCTGACCCC
 TCTAGAGATAGAGTTTTCCCTTCGGGGGACAGAGTGACAGGTGGTGCATG
 GTTGTCTGTCAGCTCGTGTCTGTAGATGTTGGGTAAAGTCCCGCAACGAGCG
 CAACCCTTAAGCTTAGTTGCCATCATTAAGTTGGGCACTCTAAGTTGACTG
 35 CCGGTGACAAACCGGAGGAAGGTGGGGATGACGTCAAATCATCATGCCCT
 TATGATTTGGGCTACACACGTGCTACAATGGACAATACAAAGGGCAGCGAA
 ACTGCGAGGTCAAGCAAATCCCATAAAGTTGTTCTCAGTTTCGGATTGTAGT
 CTGCAACTCGACTATATGAAGCTGGAATCGCTAGTAATCGTAGATCAGCAT
 GCTACGGTGAATACGTTCCCGGGTCTTGTACACACCGCCCGTCACACCACG
 40 AGAGTTTGTAAACACCCGAAGCCGGTGGAGTAACCATTGAGGCTAGCCGTC
 GAAGGTGGGACAAATGATTGGGGTGAAGTCGTAAACAGGTAGCCGTATCGG
 AAGGTGCGGCTGGATCACCTCCTTTCTAAGGATATATTCGGAACATCTTCT
 ACGAAGATGAGGGAATAACGTGACATATTGTATTCAGTTTGAATGTTTAT
 TAACATTCATTTGTACATTGAAAACTAGATAAGCAAGTAAGATTTTACC
 45 GCAAAACCGAGTGAATAGAGTTTAAATAAGCTTGAATTCATAAATAATCG
 CTAGTGTTCGAAAGAACAATCACAAGATTAATAACTAGTTTTAGCTATTTA
 TTTTGAATAACAATTCAAAATATGGTGGAAACATAGATTAAAGTTATTAAGG
 GCGCACGGTGGATGCCTTGGCACTAGAAGCCGATGAAGGACGTTACTAACG
 ACGATATGCTTTGGGTAGCTGTAAGTAAGCGTTGATCCAGAGATTTCGAA
 50 TGGGGGAACCCAGCATGAGTTATGTCATGTTATCGACATGTGAATTTATAG
 CATGTCAGAAAGGCAGACCCGAGAACTGAAACATCTTAGTACCCGGAGGAA
 GAGAAAGAAAAATCGATTCCCTGAGTAGCGGCGAGCGAAACGGGAAGAGCC
 CAAACCAACAAGCTTGCTTGTGGGGTTGTAGGACACTCTATACGGAGTTA
 CAAAAGAACATGTTAGACGAATCATCTGGAAAGATGAATCAAAGAAGGTAA
 55 TAATCCTGTAGTCGAAAACATATTCTCTCTTGAGTGGATCCTGAGTACGAC
 GGAGCACGTGAAATTCCGTCGGAATCTGGGAGGACCATCTCCTAAGGCTAA
 ATACTCTCTAGTGACCGATAGTGAACAGTACCGTGAGGGAAGGTGAAAA
 GTACCCCGGAAGGGAGTGAAAGAGAACTTGAACCGGTGCTTACAAGTA
 GTCAGAGCCCCGTTAATGGGTGATGGCGTGCCTTTGTAGAATGAACCGGCG

AGTTACGATCTGATGCAAGGTTAAGCAGCAAATGTGGAGCCGAGCGAAAAG
 CGAGTCTGAATAGGGCGTTGAGTATTTGGTCGTAGACCCGAAACCAGGTGA
 TCTACCCCTTGGTCAGGTTGAAGTTCAGGTAACACTGAATGGAGGACCGAAC
 CGACTTACGTTGAAAAGTGAGCGGATGAACTGAGGGTAGCGGAGAAATTCC
 5 AATCGAACTTGGAGATAGCTGGTTCTCTCCGAAATAGCTTTAGGGCTAGCC
 TCAAGTGATGATTTATTGGAGGTAGAGCACTGTTTGGACGAGGGGCCCCCTCT
 CGGGTTACCGAATTCAGACAAACTCCGAATGCCAATTAATTTAACTTGGGA
 GTCAGAACATGGGTGATAAGGTCCGTGTTTCGAAAGGGAAACAGCCCAGACC
 ACCAGCTAAGGTCCCAAAATATATGTTAAGTGGAAAAGGATGTGGCGTTGC
 10 CCAGACAACTAGGATGTTGGCTTAGAAGCAGCCATCATTTAAAGAGTGCGT
 AATAGCTCACTAGTCGAGTGACACTGCGCCGAAAATGTACCGGGGCTAAAC
 ATATTACCGAAGCTGTGGATTGTCTTTGGACAATGGTAGGAGAGCGTTCT
 AAGGGTGTGCAAGCATGATCGCAAGGACATGTGGAGCGCTTAGAAGTGAGA
 ATGCCGGTGTGAGTAGCGAAAGACGGGTGAGAATCCCGTCCACCGATTGAC
 15 TAAGGTTTCCAGAGGAAGGCTCGTCCGCTCTGGGTTAGTCGGGTCTTAAGC
 TGAGGCCGACAGGCGTAGGCGATGGATAACAGGTTGATATTCCTGTACCAC
 CTAGTATCGTTTAAATCGATGGGGGACGCGAGTAGGATAGGCGAAGCGTGC
 TGTTGGAGTGCACGTCCAAGCAGTAAGGCTGAGTGTTAGGCAAATCCGGCA
 CTCATAAGGCTGAGCTGTGATGGGGAGAGGAAATGTTTCCTCGAGTCGTT
 20 GATTTTCACACTGCCGAGAAAAGCCTCTAGATAGATAACAGGTGCCCGTACC
 GCAAACCGACACAGGTAGTCAAGATGAGAATTCTAAGGTGAGCGAGCGAAC
 TCTCGTTAAGGAACTCGGCAAAATGACCCCGTAACCTCGGGAGAAGGGGTG
 CTCTTTTAGGGTTCACGCCCGAAGAGCCGAGTGAATAGGCCCAAGCGACT
 GTTTATCAAAAACACAGGTCTCTGCTAAACCGTAAGGTGAT
 25

Sequence 3379

step.1002c06.cons.ok

TATCTAAAACCTCTTCATAAAGTGGACAAGAAGGTAAAGTTAAGTTATCCA
 30 TTTGAACTTTAATTAGATGTAATATTCTGTCTGCATCTTTATTCAATTGGT
 CATATGCCGCACTATTTAACTTCGATTGGTTAGACATGACATATCCCCCT
 ATCGTTTCAATTAATATCTTCACAATAGAAGTTTATCCTACTGAATGTAA
 AAAGCAAGCGAATTTAAATTAAGTGAAAGCGATTCTATGTTAAAATAATGA
 AAGATAAAAATTTAGGGAGTGGGCATTTATGATTCAAATAAAAGGAGCAGT
 35 CAATTTCCCTATTTTCATTGGATAGTACGACTTGGATATTTGACGATAGAAA
 AGTTACTATTGATGATTTAGAACGTGGGGTATTTGATGGTACTAGACCCAT
 CAACTTTGATGATGAACAAGGAATGGAACCGTGCTATTTTAGAAGGACAAAC
 AAATCCACCAACGCTTGATTTCTGAAATTAATATAAAAAACGATCGGTTTT
 AGATGAACATTTGTGAATATTTATTTTCTTTTATCTTCATAACGTTCT
 40 TTAGCTAAATCAATCAATTTAGCAATTAAATCAGGGTAAGATAAGCCATA
 TTTTTCATAAAGTTTGGATACATACTATATGCAGTAAATCCTGGCATGGCA
 TTTGTTTCATTAATATAAATTTGATTATCATCAGTAACAAAGAAATCTGCA
 CGAACTAATCCCGAACAATCAGTAGCTTTAAAGGCTCTAATGCCATGTTT
 CTTAATGTCATTTGAACATCTTGATCTAAATCTGCTGGAATATCTAATCTA
 45 ATCTTACCGTCTTTTATACTTTGATTTTATAATCATAAAACGCTACATCCTTA
 ACAACCTCACCAGGCCATGTCGTTTCAGGATAATCGTTACCTAAGACAGCT
 ACTTCGATCTCTCTAGCATTAATCCCTTGTTCAATGACAAGTTTACGATCA
 AATTGGAATGCTTCAGCTATCCCAGATTTTAATCTTCTTCAATTGTTACAT
 TTACTTTATACCAACACTTGAACCGAGATTAGCAGGTTTTACAAATACCGGA
 50 TATGTTAATCACTTAACATAATTTAATGATATATTTTTCATATTTTCA
 TACTCACTTCTTAAAAAGCTAATATAAGGTAATTGAGGTAAACCTCTATGC
 TCAAATAATGTTTTCATCACGAGTTTATCCATTGAGCTTGAAGCAGCTAAC
 ACACCATTACCTACATATGGTATATCAAGTACTTCAAAAAGACCTTGGATA
 GTTCCATCTCTCCATTTGGACCATGCAATAATGGGAATACTGCATCATAT
 55 GATTTTCTTAAACTACCTTACTGAGTAACGTGAGATTCTCTCAGTCTCT
 ACATCGTTAATGACGAGTTCATCAGTATTTTTTATTTCTTGTGTAATATTA
 TCTTTTTTCTTCCATTACCATCGTTTGTTATATAAATGATATCAACTTGA
 TATCGTTCTTTATCAATTGCGTTTAAACATTTTGTGCAGTTAAATTGAA
 ACAT

Sequence 3380

step.1002c07.cons.ok

5 TATATCGCTTTTGTGCTTAAGGTTTAAATCGTAACCATCAACATTAATGG
ATAACGTGTAATATTTTTCTGGTTTAAATAGTTAATCTCATCTTGTCTTG
ATTTAACFATTTGTATTGTTGGAGTTTGTACACGACCTAATGATAATTGTG
CATCATATTTTGTGCTTAGTGACGCTGTTGCATTAATCCCTACTATCCAAT
CTGCTTCACTTCGTGCAAGTGCTGCTTCATATAAAATTTGATACGCGTTTC
10 CATTTTTTAACGTGTTTAAATCCTTCTTGATAGGCTTTTTTTGTAACCGAAC
TAATCCACAAACGCTTGATTGGTMTTTTATTACCTACTTTATCTAAAATAA
GACGAGCTACTAGTTCACCTTCTCGTCCAGCATCTGTTGCTATAATAATTT
CTTTAACATTTTTATCTAAAAATAAGATTTTACAATTTTAAATTGCTCTAT
TTGTTTTACTAATCACTACTGTTTTTCATTTTCTTAGGAATGATAGGTAAGT
15 CTTCTAAATTCATTCTTTATATGAAGGGTTATATTGTTTCAGGCGTTGCAT
TTGTCACAAGATGCCCAATGCCCAAGTTACTATATACTGTTTCCCTTCTA
TATAACCATTACTTTTTTGTGTAAGATTCAAAGCATTAGCGATACTCTCTC
CGACAGATGGTMTTTCAGCTAAGATTAAAGATTTCAATTTAATTTTCCTTT
CACATTTATAAAATACGTCATGTTATAGAATTACATACATTGTAACATGTA
20 TCTTTTTGTAAATTACTATATATAAAATTAATTAGATAGTTATATATACA
TATCTTGTTCATCATATAGTTCAAAAACACATTATCACTTAATACGAATA
TTTACTTACTAAAAATTTCAACACTTTTAAGGTGTTGTTTTTACTTCTTT
TTTTAAATTACTATAGAATGATTTTCGAGTGTAATTGGGACACTTAATTTTGT
TCTTAGGAGAGTGAGTAATAAGTGCAAATTGTAGATTTTTTGATAGCACTT
25 TTACCAGCCCTATTTTGGGGTAGTGTAGTCATTATAAATGTTTTTGTAGGT
GGTGGACCATATAATCAGATTCGAGGTACAACTTAGGTACACTTTTATC
GGATTTTCTCTACTTGTCTACTGGACACGCAGCGTTTGATAACCTTACAGTA
ATTATTTGTCGGTTTTAGTATCAGGAGCTCTATGGGCTTTTGGTCAAGGTAAT
CAATTAATAATCAGTGCATTTAATAGGTGTATCTAAAACGATGCCTATTTCA
30 ACGGGTATGCAACTTGTTCGGTACCCTCTATTTAGCGCTATTTTCTTAGGT
GAATGGAGCAGGATTGTTCAAGTAGTGATGGGACTTATAGCAATGATCTTA
TTGGTTGTAGGTATTTCTTTAACATCACTTAAAGCCAAAAGCGAAGGCAA
TCCGATAACCCAGAATTTAAAAAAGCAATGGGAATATTACTTCTATCAACA
ATCGGTACGTTAGGTATGTCGTTCTTGGAGATATTTTGGAGTAAGTGGT
35 ACAGATGCTCTCTTCTTCCAATCAATTGGTATGGCAATTGGAGGATTAATC
CTTTCAATGAATCATAAATACTTCAATTAATCTACTGCTCTAAATCTTATA
CCAGGTGTATCTGGGGTATCGGTAACCTATTTATGTTCTATTACAAACCT
AAAGTTGGTGTAGCAACTAGTTTCTCATTATCACAACCTGCTTGTATTGTT
TCAACTTTAGGGGGTATCTTTATTCTAGGGGAGAAAAAGATCGTCGCCAA
40 ATGATTGTTATTTGGTCAGGTATTATCGTTATAGTTATAGCTTCAATCATT
TTAGGCAACTTAAATAGAATTTTAAACTTTAGGGAGGTAACATAGTGT
GAAGAATTAGAAAATAAAGTGGTCTTATTACTGGAGCTGCCACTGGAATT
GGCAATCTATTGCGGAAAATTTTGGTAAAGCTAAGGCCAAGTTGTTATA
AATTACCGTTCTGATCGACATCATGATGAAATTGAGGAAATTAACAAACT
45 GTTGTCTAAATTTGGTGGTCAAAACATTGGTGGTTCAAGGTGATGTTCAATT
GAAGAAGATATTAAACGAATGATTGAAACAACAATTAATCACTTTGGAAC
TTAGACATATAATTAATAATGCTGGATTTCGAAAATTCAATCCCAACTCAT
GAAATGTCGATTGACGACTGGCAAAAAGTTATTGACATAAACTTAACTGGC
GCCTTTGTGGGTTCAGAGAAGCCATCAATCAATTTTAAAGGAAAAACAAG
50 AAAGGTACTATTATTAAACATTTTCGAGTGTTCATGACACTATTCATGGCCT
AATTATGTACACTATTGCGCAAGTAAAGGTGGCTTAAATTAATGATGGAA
ACAATGTCAATGGAATATGCCCAATACGGTATTCGTATTAATAATATATCT
CCTGGGGCAATTGTTACTGAACACACTGAAGAAAAATTTCTGACCCAACG
ACGCGTGAAGAAACAATAAAAATGATACCTGCACGTGAAATTGGAAATGCT
55 CAAGATGTAGCTAATGCAGTACTATTCTTATCTTCAAGTCTTGCAAGTTAT
ATACACGGTACAACATTTGACGTTGATGGTGGCATGATGAACATATCCAGCA
TTTATGGGTGGTAAAGGTAAAGTAGCATTTTCAATTTTAACTAGAAATAA
AATGTATATAAATTAGTAATGAATAGTAGTCTTAAAAAGTATCAGTCATCGA
CTGAATATAAGTGACAACTATAATAAAAAAGCTATGAAGATATCTACAAT

5 GATACATTTCTTCATAGCTTTTCTTAGTTATATAAATAAGTATATTGATTAT
GCAATCAATTTTTATACAAATACAGCTTGATAAATATACAAAATGATAATG
CTAATATGAATATTTAGCAAACAATTCCGTTTCTAATGATGATCAATTATC
ATCTGAAGAACAGCAGAAAATACAATGAAAATACTAACGAAAACAAACAGAA
TAGTCATGTTGACTTAAATTCATTACCTGCTACAGATTTTCGAACAGATTG
10 GATGCTGAATCTGGTCAAAAACAAGTGAATGATTAGCAAACCAAAAAGA
TAGCGGTCAAATTTACAAATACGAATTTAATGATCGCGTATCTGATGTGAT
GAATAATGAAATAGATAATCAACAGCCATAATTGAAATATAGGGTAGCAGC
TCTACCCTTTTATTTAGGGAGGGTAAACACGTGGCATAAGAAATTCACCT
ATAAACATGGTGAAATCAAAAATCGCTATTATGAGAAGTACAAAGACCCTC
15 TCACAAACAAAATGACGACGTGTTAGCGTGGTACTTAATAAAGAATGGTAAA
CAGTCACAAAAGAGGCTCAGAAGCGCTTAAATGAGCTTATAGAGGCAAAA
TTAAACGATAAGACACCTACTACACTTAAGTCACTAACATTCCATGATCAT
TTGATGAGTGGTTTTGAAAGATACAAATTAAC'TCCGGATTTAAACAACTA
CCGTTACAACGAAAAGTTATAAAATAGCGCACATCAAAAAGAAATATTGATA
20 AAAATATACTCGCTCAAAATACGAATGCTGATGTTTTACAAGACTTAATTA
ATTC'TTCATTAAACAGAGGGATTAAGTCATAAAGTGGTTAAAGATGATTGA
GTAT'TATAAAAAATATACTACGCTATACTCAAAAGAAATATAACAT'TACTG
ATAAATCTTACATAGATGATGTAATAGTACC'TAAAAAGCGACTACAATAG
AAGAAGTTAAAGCAAAACGTGAGAATATTTAGAAATGAATGAAGTTCATC
TTATCGCTGAAGAGTT

Sequence 3381
25 step.1002c12.cons.ok
GCTTTTCTCAAATTATCTTTAACACGTTCCCTCATTGACTTTAGTTTGATGT
TGTTCTCTTTTAGCTTGATT'TTCCTTCTCTATTTTCTCACGCTTATGCGC
TCTTCTCGCTCTACGTTATGAAAATATTCACGCTCTTTACGTCGATATTGA
30 CTGCGTGGTATGACTTGT'TTTTATTGTTATCCACTAGAAAGACCACCTCCTA
TAGTTTCTTTAACTCATCTTTATCTGTAAATTCAGTTATTTTAACTACT
ATATATCATCATCATTTTAAATTCGTTTATAATCTAAAGTTTAAATATTTTA
ACTATAAAACTGCAATGTTCTAATAAAAAAT'AAAAATAAAAAGAACCTCCAC
TTAAAT'TATTATTAAGCAAAGGTTCTAGCATATCAATCAAAT'TATTTTGA
TTACCT'TTTTCTTGTGTGCAAT'TGCTGAAGTTAACCATCCTAT'TAAACT
35 AAGCCAACTAATACTACCCAGAAGATTGTTTGCCATAAAGCACTATGTGGA
AATGCTTCTGGTAAAACGCCAATATCAGGATGTGCAAGTACCATTATAACA
AGTTTAAATACCTACCCAACTTACAATTGCAAACGAGCACC'TCAAGTCTC
GGATATTTATTCAACAAT'TCTACAAACCAAGTTGCTGCAAATCTCATCAAG
ATGACACCTATCAT'TCCACCAAGGAACATAACAATAAAT'TGGCCTAAGTCC
40 ATACCACCAAAATGTATGCCAACTTTTGGTAAATGTAACGGCTATGGCTAAT
GCGGCAAGCATCGAATCAATTGCAAACGCGATATCAGCGAAT'TCAACTTTA
AATACTGTTCCCAAAAAGATTTAGGGCCTACTTCTTTTTCGTTGCTGTT
TCATCGAAATGATGTTCTGCTCCTGTTCTTTATGGTGT'TTTTCATTTGAC
TGATGGAAAAAT'TGCCATAAATTTTAAATAGACATATAGATTAAAGTAAACA
45 GCACCTGCTGCTTGAATCCACCAGAAGTTTGCAATAATACTTATTAATAAT
AAAGCAATAAATCTAAAAATGAATGCACCTAATAGGCCATAAAAAAGTGCT
TTTTTACGTTGTTT'AGGTGGTAGATGTTTAAACCATTACCGCATTACAATT
GCATTATCAGCTGCTAATAATCCTTCTAAAAATACAAGTACT'AAAAGTACC
CATAAATAAGGTAATACTCAAACTCGGATCCATTGTGGACAAC'TCC'TTTAG
50 TTTAAATACAAATAAAGAGACCTATGCTAATTTAATAGCAAAGGTCTCACT
TAGCAATATTTTGGCCACTTTACCGGAAGACTTATCTTCGTAAATGACGAC
AAAGTGT'TGAATAGTTAT'TATCGGTTTAAATGCATTCAATAACTATTCTCA
AGTTACTCCCTCTGACAGTTTGTGTCATAGGTATTTATTTGTTATTTTTCAG
TATACATACTATGACACTGCATAGCAATATAAACATTTTTCATATACTATG
55 GTGTGCCCGAATAGTGATATGTTTAAACATTC'AATTATAATAAATATACAA
GTTTAAATATCAGAGAATTTTCTTCAAACCAACGCGTTGCAAATTCATTTT
CAAATAGGAATACATAAT'TGTCGTATCTATCTTTTACTAAAATGATCTAG
AAGTATTCATATTATCTTTAATATCTTCTTCATTTTCAATCCAACGCGCAA
TTTACGTCCAAC'TGGTTCCATAACCACATCTACATTATATTCATTATTCA

TGGCGGTGTTCAAACACCTCAAATTGTAATTGACCCACAGCACCTAAGATAA
 TTTGATTTGTGTGCAACGTTTTATAATATTGAATTGCACCTTCTTGAACAA
 GCTGTTCAATCCCTTTATGAAAATGCTTTTGTTCATTACATTTTGGCAG
 AAACCTTCATAAATATTTCTGGAGTAAATTGTGGTAACTCTTGGAACTAT
 5 ATTTCTGCTTACCACCAACTAAAGTATCGCCTATTGATAGTTACCTGTAT
 CATACAAACCAATAATATCGCCTGCTACAGCATGATTAACAGTTTCTTTAT
 CATCGGCCATAAATGATGTGGAACGAGTTATTTTGTGTTTTATTTAGTTC
 TTTGTAAGGTTACATCCATTCCTCTTTCAAACGCACCACTTACCACGCGCA
 TAAATGCAATTCTGTACAGATGCTTTGGATCCATATTGGCTTGAATTTTAA
 10 AGATAAACCAGAAAAATCTGTGTCGAATGGACTAACATCAACCTCTTCTT
 TAGTTTGTCTCGCATTAGGCATTGGCGCATGATCAACGTATGCATTTAAGA
 AATTTTGTACGCCGAAGTTTGTCTAATGCCGATCCAAAGAAAACAGGTGTTA
 ATTCACCGTTCAGTAACGCTTCATTATCAAATGCTTCACCCGCTTCATCGA
 CAAGCATCATTTCTTCAATGCTTGTTCAAATGCACTATCATTTTATTCG
 15 CGTGTCTCTCTTTAAGTTCGTAATCTTCGTTTAAATGTAATAAATTTCTT
 CATCTCTAAATGGTTCAATCGTTTATAGAGTGTCTATCAATAATACCAAAGA
 AATTTTGTCCCATACCAACAGGCCAATTCATTGGATAAGTGTCAATATTAA
 GTGTTTCTCTATTTTCATCAAGTAATTCAAATGGCTCTTTACCTACTCTAT
 CCAATTTATTGATGAATGTAAAAATAGGGATACCTCGCATTTTACAACTT
 20 TAAATAATTTAATGTTTGTGGTCTATACCTTTTGCACAGTCAATAACCA
 TAACGCACTATCTACTGCCATTAATGTACGATACGTATCCTCTGAAAAAT
 CTTTCATGCCCCGGCGTATCTAAAAATATTTATCTTATAATCATCATAATCAA
 ATTGCATACAGAGCTTGTAAGTGTGAGATACCAGTTCCTTGTCTACCTTCA
 TCCAATCACTTGTGCAAATTTACCCGCTCTTTTCCCTTTAACTGTCCCAG
 25 CTTACGAATCGCTCCACTGAAATACAACAATTTTTCAGTTAAAGTGTGTT
 TACCAGCATCAGGGTGTGAAATGATGGCAAATGCTTTCTTGACTCTATTT
 CTTCTTTTAAATTCATTTATCATTTCTCCTTTGGATTCAATTTCTAATCAAAT
 CAGTTATATTACGTGCCAATGCATCTGCTTCAACTTCAGTCAACAAATTTGA
 ATAATTGCATCACTAATTCCTCATATAGCTGATTGTCATCTAATCTTTGG
 30 ATAACCTCAATGGACCAAAAAAGTTCTGGAATAGCTAACATGATTAAATCAT
 TAGAATGATAATGATATACGTTTCATCATCATTCATGAAACTGAAGTGATT
 TAATCCTCAAGAGGTCCACCACCAAAATTTTGTATGCTGCTTTTAAGCCA
 ATTAAGTCATCGCGATGTGGGACTTTAACATGACCAGGCATAATTTGATAA
 GGCTCTCGACCTTTTGAGGCCAAAACAACGTATCACCTGGTTCAGCAATA
 35 TCAATCGCGTGTCTAATACCTTCTGCACGGTCATCAAACCTCTATATAATTG
 TTATGCGTTGCACCTTTAGCTAATTCAGCTGTCAACATTTTAGGATCATCG
 TTAGCAGGATTATCTGGAGTAAAAATAACATAATCTGCTCGACATGCTACT
 CTCCCCATCTCAGGTGTTTTCTGTTAAATCTCTTTCCCGCCATACCTACT
 AAAAAGATTAGCTTTTGTTTAACAATGGTTGTACAGCATCAATCAGTTTA
 40 TTCATACCATCAGCTGTATGTGCATAATCTATAATCAAATCTATAGGAAGT
 GATGGATCAAGTACTTCTAAGCGCCCTCTACAGGTTCAAGCTCAGTAACA
 GCATTAATAATTTCAATTTAAATTTGTTCCCTTACTCCAAACCGCAATCATA
 GCAGCCATAATATTTGAAATGTTAAACCTACCTACATAAGGAGATTAACT
 GGAAACTTCCAAATGGAGTACAAAATTCAAATCTACACCTTGCAATGAT
 45 TCTTTAATGTTTTTAGCCATAAATTTGTGCTTCGTGAGTAATACCATATGTA
 AATACTTCATAGGGTGTACACTTGCTAAATAATCGGAGAAG

Sequence 3382

50 step.1002d01.cons.ok

CTGTCGTAAAGCTGGTATTATTACTGGTTTGCCAGATGCTTACGGACGTGG
 ACGTATTATCGGAGACTATCGTCGTGTGCTTTATACGGTGTAGATTTCTT
 AATGGAACAAAACTTAAAGACTTTAACACAATGTCTACTGAAATGTCTGA
 AGATGTAATTCGTTTACGTGAAGAATATCAGAGCAATATCGTTCACTTCA
 55 AGATTTAAAGAATTAGGTCAAAAATATGGATTTGATATTAGCCGTCCTGC
 TACTAACTTCAAAGAAGCTGTGCACTGAGAGATCCCTCATAATTTCCCA
 AAGCGTAACCATGTGTGAATAAATTTTGAGCTAGTAGGGTTGCAGCCACGA
 GTAAGTCTTCCCTTGTTATTGTGTAGCCAGAATGCCGCAAACTTCCATGC
 CTAAGCGAACTGTTGAGAGTACGTTTCGATTTCTGACTGTGTTAGCCTGGA

AGTGCTTGTCCTCAACCTTGTTTCTGAGCATGAACGCCCCGAAGCCAACATG
 TTAGTTGAAGCATCAGGGCGATTAGCAGCATGATATCAAAACGCTCTGAGC
 TGCTCGTTCCGGCTATGGCGTAGGCCTAGTCCGTAGGCAGGACTTTTCAAGT
 CTCGGAAGGTTTCTTCAATCTGCATTTCGCTTCGAATAGATATTAACAAGTT
 5 GTTTGGGTGTTTCAATTTCAACAGGTAAGTTAGTTGCTAGAATCCATGGCT
 CCTTTGCCGACGCTGAGTAGATTTTAGGTGACGGGTGGTGACAATGAGTCC
 GTGTGAGCGCTGATTTTTTCGGCCTTTAGAGCGAGATTTATACAATAGAA
 TTTGGCATGAGATTGGATTGCTTTTAGTCAGCCTCTTATAGCCTAAAGTCT
 TTGAGTGACTAGATGACATATCATGTAAGTTGCTGATAGGTTTCCAGTTTT
 10 CCGCTCCTAGGTCTGCATATTGTACTTTTCCCTCTTACTCGACTTAACCAGT
 ACCAACCCAGCTTCTCAACGGATTATACCATGGCACTTTAAAGCCAGCAT
 CACTGACAATGAGCGGTGTGGTGTACTCGGTAGAATGCTCGCAAGGTCGG
 CTAGAAATTTGGTCATGAGCTTTCTTTGAACATTGCTCTGAAAGCGGGAACG
 CTTTCTCATAAAGAGTAACAGAACGACCGTGTAGTGGGACTGAAGCTCGCA
 15 ATACCATAAGCCGTTTTTGTCTCACGGATATCAGACCAGTCAACAAGTACAA
 TGGGCATCGTATTGCCCCGAACAGATAAAGCTAGCATGCCAACGGTATACAG
 CGAGTCGCTCTTTGTGGAGGTGACGATTACCTAACCAATCGGTGATTCGTT
 TGATGTTATGTTTTGTTCTCGCTTTGGTTGGCAGGTTACGGCCAAGTTCGG
 TAAGAGTGAGAGTTTTACAGTCAAGTAAGGCGTGGCAAGCCAACGTTAAGC
 20 TGTTGAGTCGTTTTAAGTGTAATTCCGGGCGAGAAATGGTAAAGAGAGTCGT
 GTAAATATCGAGTTTCGCACATTTTGTTGTCTGATTATTGATTTTTGGCGA
 AACCATTTGATCATATGACAAGATGTGTATCTACCTTAACCTTAATGATTTT
 GATAAAATCATTAGGGGATTCATCAGAGCTGTGCAATGGTTATACTTAGC
 ATATTTAGCTGCTATCAAAGAACAATAATGGTGCAGCAATGAGTTTAGGACG
 25 TACTTCAACATTTCTTAGATATTTATGCTGAACGTGATTTACAAAATGGTGA
 CATCACTGAACAAGAAGTTCAAGAAATCATTTGACCACTTCATTATGAAAT
 GCGTATCGTTAAATTCGCGCGTACGCCTGAATAATGAATTATCTCTGG
 AGATCCAACCTTGGGTAACGAATCTATCGGTGGTGTAGGTATTGACGGACG
 TCCAATGGTAACAAAACTCATTCCGTTTCTTACACTCATTAGATAATTT
 30 AGGTCAGCAGCAGCAACCAAACTTAACAGTGTTATGGTCTACTCGCTTACC
 TGAAAACCTTCAAAATCTATTGTGCTAAAATGAGTATTAACGAGCTCAAT
 CCAATATGAAAATGATGACTTAATGCGTGAAAGCTATGGCGATGATTATGG
 TATCGCTTGTGTGTATCTGCCATGAAGATTGGTAAACAAATGCAATTCTT
 CGGTGCACGTGCTAACTTAGCTAAAGCATTACTTTACGCTATCAATGGTGG
 35 TAAAGATGAAAAATCTGGTAAACAAGTTGGGCCAAGTTATGAAGGTATTAA
 ATCAGACGTACTAGATTATGATGAAGTCTTCGAAAGATATGAAAAATGAT
 GGACTGGTTAGCTGGCGTATATATCAACTCATTAAATATCATTCACTATAT
 GCATGATAAATATAGCTATGAACGTCTTGAAATGGCTTTACATGATACAGA
 AATTATTCGCACAATGGCAACTGGTATTGCCGGATTGTCTGTAGCAGCTGA
 40 CTCTTTATCAGCGATTAAATATGCACAAGTTAAACCTATCCGTAAACGAAGA
 AGGTCTTGTAACGACTTTAAAAATCGAAGGCGACTTCCCTAAATATGGTAA
 TAATGACAGTCGTGTTGATGAAATTGCAGTAGATTTAGTTGAACGTTTCAT
 GACTAAATTACGTAGCCATAAAACATACCGTAATTCTGAACACACAATGAG
 TGTATTAACAATTACTTCAAACGTTGTTTATGGTAAGAAAACCTGGTAACAC
 45 ACCAGATGGACGTAAAGCTGGCGAACCATTTCACCTGGCGCAAAACCAAT
 GCATGGTCGTGACCAAAAAGGTGCATTATCTTCACTAAGTTTCAGTAGCTAA
 AATACCTTACGATTGCTGTAAAGATGGTATCTCAAATACATTTAGTATCGT
 ACCGAAATCACTAGGTAAAGAAGAAGCAGATCAAAATAAAACTTAACCTAG
 TATGTTAGATGGTTATGCAATGCAACATGGTCATCACCTCAACATTAACGT
 50 ATTTAATAGAGAAACATTAATTGATGCAATGGAACATCCAGAAGAGTATCC
 ACAATTAACCAATTCTGTTATCTGGATACGCTGTAAACCTTCAATTAATTAAC
 ACGTGAACAACAATTAGATGTTATTTACGTACATTCCACGAATCTATGTA
 ATAAATTTAAGGTGGGAGCAAAATGCTTAAAGGACACTTACACTCCGTTGA
 AAGTATGGGCACTGTGACGGGCCAGGACTTAGATATATATTTACTCA
 55 AGGTTGTTTGTAAAGATGTTATATTGTCATAATCCAGACACTTGAAGAT
 TAACGAACCATCAAGAGAAGTGACGGTTGATGAAATGGTAAATGAAATCTT
 ACCGTACAAACCTTACTTTGAAGCTTCAGGTGGTGGGGTAACAGTCAGTGG
 TGGCGAACCATTACTACAAATGCCTTTCTTGGAGCAATTATTCAAAGAATT
 AAAAGCGAATGGTGTTCACACATGCATTGATACTTCTGCGGGATGTGTGAA

TGATACACCAGCATTTAATCGTCATTTTGATGAATTGCAAAAGCATACAGA
 TTTAAACATATTGATAATGATAAGCACATCAAATTAACAGGCAAACCTAAC
 ACACATATTTTAAAGTTTGCACGTAAATTATCTGATATGAAACAACCTGTT
 TGGATTAGACATGTTTTAGTACCTGGTATTTCCGGATGATAAAGAAGATTTG
 5 AATAAACTAGGAGAATTTATTAATTCCTTTAGATAACGTTGAAAAGTTTGAA
 ATCTTACCATATCATCAACTCGGTGTGCATAAGTGGAAAAATTTAGGCATC
 CCTTATCAACTCGAAAAATGTTGAACCATCTGACGATGAAGCGGTTAAAGAA
 GCTTATCGCTATGTTAACTTTAATGGCAAAATACCCGTAACATTATAGAAA
 GATAAATATAGTAATAGCCTGAGACATCAATCGTTGTCTCAGGCTTCTTTA
 10 CGTTTAAATCAGTCGATAAACTGAGTATAAAATATATCTCAATTAACTTTAT
 TTAATAAAAAATATGATTTCATAAGTTTAAATAGTTTATTTTCTAGATATGC
 TTTAAATACATCTTTATAAACATCGATGTAACCTAAATACATATCTTTTCT
 AATAAATTCATCAACTTGATGCGCCATAATAGATTCACCAGGTCCAAAGAC
 AGCAAAATCCACGTTATTTTCATTTGTTTCTAGGAAACTAGATGCATCAGT
 15 TGTACCAATCAATGCACCTCACAACAATGTCTTCATGTACATAATTCGGTGC
 AATACGTGTAATATTTTGAATAAGAGGATTATCAGATCACTTGCCACTGG
 ATCGTGACTGTAGGTATATCTACAGTTAAATAATCTTCGCCACATGACG
 AATGACTTTTTCAAATAAATCCTTCACGAAAGTACTGTCATATTTCTGGAAC
 AGTTCTTACATTATATTTAGCTGTCTGCTTTATGAGGAACAGAATTAACCTG
 20 TTTACCACCATTTGAATACAGAGTTTAACATTACAAATCCAGAGTAGATATG
 TGATTCTTCTTCACCAATTTTCTGTGGAGATGTTTCTCAATCATTTGGAAC
 AGCGTCTAATCGTGTACTTTATCATGTTCTTTAATATTTTATATTTCTTG
 TTTTCAATTTTATCAAAAATCAACTAAAATATCAACAGCATTTGTACCTAG
 GTGTGGCATGGAACATATGTGCAGCCTTACCTTTCGCTGTGACTACACAGCT
 25 CATAGAACCCTTTATGAGCATAGTAAGCGATATTACTAGTTGGTTCGCCAAT
 AATTAACCACCTCACATCATCTAAGTAACCTTCATCAGCTAGTAATTGTGC
 ACCATATTGTT

30 Sequence 3383

step.1002d02.cons.ok

GAGATAATCATAAAATTAGATACTACAGTCATAGTCAGATTAATTTGTCTT
 AGGACTAATGCTTTTTTCTTTTCGGTGATAAGAGAGGACATTCTCATATTC
 AAAAGGTGTTTTCATCCATCTTATAAAATCAACAAAGTTATCAAAACTAAC
 35 TTCTTTATAGCATATCTTTTTACTTTCTAAGTAAACAAAAAATTTTTTAG
 TGCATAGGCATAGGTTTTCTTGGTATTTAAACTTTTCTTAACACTATCCAG
 ATACTTCAAATATCTTTACTGCACTACTATAGGTTTCAATTATTACCATCTAA
 AATCATAAAATTTGGTACCATTCTTAGATTTTACTTCTACTATTTTCATAAT
 AATCTCATTTCTAATTTTAAATAGTATAAGTATACACTATTCAAATTCCTTT
 40 TACATCATACATATTAACTTATTACACGTAATAAAAAAGACTATTTAAT
 ATTTTTACTTTGTTCAATTTTCTTAAATATGATTCAATTTTAAACAAAT
 AGTCTATATTTCTTCAACATATATATAACCATTTCGCAACTTTACCAGAACAG
 TTTGAAACAATTTATCAATATATGATAGATCAAAATAAAGTGAGCAGACCA
 GTGCTTTCAGACGACCAACTTTCGCAACTAAATATACATTTACATGAAGCA
 45 CTACAACAATCACGCCCAGTCAATATCAAATATTACGAAGAAGGATATATT
 AATTTTATAGAGCTCATTGTACATCGGATTGATTCAATAAATTATGAGATT
 GAAGGTACTGCTCCTCATTCAAGAGAACGTCATAAAGTTTCATTTTTAGAT
 ATTATAGATATTTCTTTTATATAAAGAGTACAAGAGAATTTATTATGAATT
 TCTCATTTTACACTTATATAACCACCAACATACCTTAGTGTTATCTAAAGCC
 50 TTTCTGAATGAAATCATTTAATTATTAATTTTGAATTCATAAATTAATTT
 TGACTGATAATTGGTAAATCAACTAATAAAAAATAAGCCTACACATCTCAA
 TGATGTACAGGCTTATTTCCTAAATATTATAAATTAATTGTTGTTTTCGAT
 ATCAGAAACGATTTCTTTTCGTTTTCTTTTCGATATCTTCAAAGTCTTTTTT
 AAAAATATAAGCAAAGGCAACTACACCGATGCCTACTACTGTAGTAAT
 55 AGCTAATAAATAAAGACAAATTTGAATAACCCTTTAATAAAGTTCCACAT
 TAGTTTTACCTCTATAATTAATATGCATAGCTACATTTTCAATATGATAAC
 ATACTTAACTTTATCATTAATAATATATACCTTTTTACAGGTAATTAATCTA
 AATATGAATATAAATTAATGATAATATATCAACTAAAGATGCTTAAATGAA
 ACAATTAATGATTTCTTTAATCATGGTACACACACAGATTAACCTTAGCAA

CCAAGATTTTAAATTTATCAGTGATTACTTTGTAATTATCAATATTCAGAAT
 GATCACTCCTATTGAGATAAATGTCATTATCTTAATTTAGATTGTATGATTT
 ACGCAAAATCAAGTTCAATTAATAATAAGTTACTTTTATATTACATTGAAGT
 AAAAAATATAAGGAGAGATTATAAAATGCAAAAAGTACGTTCTGATATAATG
 5 ACTCATCGTGGTTCTCACTATGATTTAGGAGTAAAGACTGCTCTTTGGTTA
 CAAACCACCTCCTTTATTAAAAATCGAAATAAAGAATGGCGAAAGAGAATT
 CCACGTTTGTATATTGATGTCAAGGAAACCTATGATATATTTCAAATCTAT
 TCGCCACAGATTTGGGAAGAAATAATTGGTATGCAAGATGATTGAATCTA
 CCTACAAAACAAATGATTTTAAATTTTGGCCATTATCGATTTACTGATTTA
 10 AAGGACAGTGGTTGCACAGTATATAAAGGTCGTGATTTTGTAGTCCGAAAT
 TATGATTATCATCTGCAACATATGATGGTAGATACTTATTATTTCAACCT
 AATGACGGGGGATTATCTCAAAATAGGACCGACTTCAAGAGTGACTGGTAGA
 ATGGATGTGATGAACGAGTATGGTTAGTTATGGCATATAATTTTATGCAT
 CGTAAAAAGCCTGCAAAATGGATTTGTATGTTACATGGTTGGTCCGGCTAATA
 15 CTTGAAAAATGCAAAAATGTAAGTGAAGCAATCAAAATTTTAAAGGAAGTA
 CCGCATCGTAGTTTCATTTCAGTTATATACTAATGGATAGACATTGCAATTAT
 GCCATTGTGCAAGTTACACCTCGATCAATAGATGTAAGGTATGAACATATA
 TGCACAAATCATTTTGAATTGCTTACCCATGAAAAATAGAACTATACAAGA
 GAATCTAAAGAACGCTTAAATCGTGAATAAATAAAACAACCTCCTTCTACA
 20 AACAAAGATATCGCATTCAAAATTTTAAACGACCCGCAATACGAAATCTAT
 AGCAACCTATTTAAAGTTGGTCTGGTACAATTCATACTTCACATATGAA
 CCTAATTCATTAATATCATGGATGGCATTAGGTCAAAACAGTCATCCAACC
 TCAATCAATTTTCTAATTGGTTAAAAGGAAAGAAATGGAATATAAATTTAC
 TTTGAAGGCGAAATAGATACACCATTAACTTTTGGCCACATACTAATTAATA
 25 ATTGACCCCTATCTTGATAAATCAAAGGGTCGAATAATTTATATTATTTAT
 TCAATTTTAAATATCGCCATCTACAGTGTAATAATTTCAATATATTATCGCTA
 TTTCCCACTTTTCTTTTATCGAAATATCGATTTTAAATTTCTTTATTTCCA
 TGACCAGGATGAAGTTTAAAGCAAAGTGTTTTTAGGTTTTTCGCCAAAATGA
 TAATTAATATTACCTCTTTGTGTAGAAGCTTTTATATCATTTGTTAGATTTT
 30 ATGTCCGTATAGAAATATCTCTCTATTAAACCAAGAAAACAGTATAGCA
 CTCTCGTCAATAATATTAACAGCATGTGGACAAGAAAATGTAAAAAAAAC
 AGAAAGCAATCCAAAAGAAGAACAAAATA

35 Sequence 3384
 step.1002d03.cons.ok
 TCTGCACCATATTTTACTGCTTGGCTCACTTCGTAATCTGCTGCATCCATA
 ATTTTCAAGTCAGCTAATACTTTAGCATTATTAATATTTTCATTTAAATGT
 TGAAGTCAGGTAACCCCTTCATTAATTACAATTGGCGTACCAATTTCAACA
 40 ATATCTACATATTCTTCAACTTTTGTAGCTAATTTTGCTGCTTCTCTTTA
 TTTAATAAATCAATGGCTAATTTGTAGTTCCAATGAAATTCATCCTTTCTGTT
 TTGGGAAATCTCTTTTGTATAGTTGTATACCTTCGCTAACTTATTA
 AACATATAAAATATTTTGTATATAAAACAACATATAATTATAAAACAATCT
 GGCGCTGCTTCATCGTCTACATAAACTTCAACATTAGGGTGTGTGTGTA
 45 ATGGTTCGAGGTACATCTTCGGTAACCTGTTCAATTAACAGTTTACTTATA
 GCCTCTTTTCTTTTGGACCAAATGCGAGTAGGATAATCCTTTTGTCTTT
 AAAATACTTTTACCCCCATTGAAACTGCTTGTCTAGGAACATCCTTTTCA
 TTGTCAAAAAATCGACTATTTGCTTTTATGGTGCCTTCTGTTAAGTTCACC
 ACATGTGTTTCACTATTGAAGTCAGTCCCTGGTTCAATTAACCAATGTGA
 50 CCATTTTCTCCAATTTTGAATATCGATTGGCCCTCTTTCATCT
 AATAAATTTGTATATCGTTCCGCTTCTGCTTCTAGATTTTCTGATAAGCCA
 TCAGGAATATGAATGTGATTTTAAACAAAATGAGGATATTGCTCAAAATAGC
 ACTTTATTCATGTAGGTATGGTAACCTTTGTTGATGACTTGCTTTTAAACCT
 ACATATTCAATTTAAATTAATGTTCACCTGCGACACATCAACTTTATTT
 55 TTGATTAATAAATTTACTAAGTAATGATATACGTCAGTCATAGTTCCACCA
 GTTGCTAAACCGAGTTTGGCATGAGGGTGTGCTGTATTGTTTAAACAAC
 TCGCATGTACATAAAATGACGCAAGATTTTGTGAATCTAAATTAATAATT
 TTCAATATCAAAACACCTCTAATTTCAATATCACAATGATTGATGACTAACG
 CCTTACTCTTTTATATCTCTTAAATTTACTGTAAACAAATGATTAAAGATTT

TAAACATATTAACGCTCTATATGTAATATTCAACGTTATTTCTTGAATTTT
 TCCTAATTAAATCGAGCATTCGTGTACAAATATACATAGCTTTCTAGTAAA
 ATAGATAGATAAACGAGTTTGTCTGGTAAATAATATGTTTATCGGTAGTCCG
 TCACTGACATATTGTTTAGTTTAAAGGGAGTGAACGTGTTGGAACCAATTAA
 5 AGAACAAGAAAGTGCTAGATTTATTAACTTCTTACTCAAATCAGCCTGTTTA
 CCTACACGTTGAAACAACAAATGGTGCTTATGCAAATCATTTCGATCAACG
 CGTATTTAACGCTGGAACATTTTAAAGAAATATTGTCTGACTTTTGAACA
 TGCACAACTTAAAGGCGGCGACAAAGATCCATATCGGTAGGTCTTAAATT
 AAAAGATGGTGGCTGGGTTTACGTGCAAGGACTTACGCACTATGAAGTTAA
 10 TGAGAATAACGAATTTTAAATTGCAGGTTTAAATTATGAAGGACAATTGGC
 TGCTACAATAGAAATAAGTAAACAGCCATTTACTATATAAAGGAGGTGCC
 AATCATGACTGATGAAAGACACGTACTTGTGATTTTCCCCCATCCTGATGA
 TGAAACTTTTTCGTCTGCTGGAACATCGCAAGTTATATTGAAAAAGGTAT
 TCCCGTCACATATGCATGTCTTACCCTAGGACAAATGGGACGTAATCTAGG
 15 TAACCCCTCCTTTTGCACAAGAGAACTTTTACCATTATACGTGAACGTGA
 GTTAGAAGAAGCATGCAAAGCAATTGGGATTACAGATTTAAGGAAAAATGGG
 GTTAAGAGATAAAACTGTTGAATTTGAACCTTACGATCAAATGGATCAAAT
 GATTCAATCATTATTGACGAAACAATCCATCATTAATTATTTTCGTCTTA
 TCCTAAATTTGCAGTTTACCCGTATCAGGAGCACTGCAGAAGCTGTAGT
 20 ACGTACAGTTGGACGCATGCATGAATCAGATCGACCCCGTCTTACACTTGT
 AGCGTTTGTAGCAATGATGCATCAGAAATCTTGGAGAACCTGATATTCAAAA
 TGACATATCTCAATATAGTGATATAAAACTTAAAGCTTTTGAAGCACATGC
 TTCACAAACAGGACCATTTTAAACAACCTTGCTAGTCCCGAAATAGATGG
 TCAAGCACAAAGTTTCTTAAAAATAGAGCCATTTTGGACATATCACTTTGA
 25 ATCTTAAATGGAGGTAAACATGACAGAATTTGACTTATCCACTAGAGAGG
 GTCGTTGGAAACATTTTCGGTTCTGTTGACCCGTGTCAAAGGTACGAAACCAA
 CTACTAAAAATGAAATGACCGATTTACAAAGTACTCATAAAAAATTTCTTAT
 TTGAAATAGAGGAAGTAGGCATTTAAAAATTTAACTTATCCAGTTTAAATTG
 ATCAGTACCAAAACAGCTGGTTTATTAGTTTTTCAACGAGTTTAAATAAAAA
 30 ATGAAAAAGGCATTAATATGAGTCGCATATTAGAAAGCGTTGAAAAACATT
 ATGAATATGGCAATTGAACCTGAAATTTAACACACTACATCAATTGTTGCGTA
 CTCTACAAGATAAAATGAATCAAATGCTGCAGGTGTTGATGTGTACAGGTA
 AATGGTTCTTTGATCGTTATAGTCCTGTGACTAATATTAAAGCTGTAGGCC
 ACGCAGATGTTACTTATGGTTTGGCTATTGAAAATCATACCGTTACACGCA
 35 AAGAATTAACATATTCAAGCCAAAGTAACAACACTATGTCTTGCTCAAAAG
 AAATTAGTGAATATTCCGCACATAATCAACGTGGTATCGTTACAGTTAAAG
 CATATTTAGATAAAAAATAATGATGTATCGATGATTATAAAGATAAAATTT
 TAGACGCTATGGAAGCCAACGCTAGTTCTATCTTATATCCAATCTTAAAAAC
 GTCCTGATGAAAAACGAGTAACAGAACGCGCTTATGAAAATCCACGATTTG
 40 TTGAAGATTTAATTCGACTAATTGCAGCTGACTTAGTTGAATTTGACTGGA
 TTGAAGGGTTTCGATATTGAATGTCGTAATGAAGAGTCTATCCACCAGCATG
 ATGCTTTTCGCACGTCTGAAATATCGCAAATAAAACATTGAATTTCAACTAC
 AATGCCAGTTGATTTTATGTTTATAAGATGCAACTGGCAATTTTTTAGCA
 TCATCAATATCGATGATACATAATGTAACACAAAAGAACTCACCCCATGC
 45 TAAACATTCATGTTGAATTTATTAAAGGTCGTTACGAATCTAAATAAATC
 AGAAATTCAGATTTAATTTATTTAGTGTATTTATGCACTTGTAATTTAC
 ATACATATTATAACGTTTACAACCAACTTTAAGTCATCAGATTAAAGTATT
 TATTAATCACACATACGTGTTTACTAACAATGAACTTCAAACAAATGCAT
 CAATTCATTTATAGGTTACCTGAGCTATTGACAACTATTTACATTTAATA
 50 CATTTCGACTAGTGAACCAAGAATGTGTCATGCTATGAAGAAAGGTGAAATT
 CATGATTACTG

Sequence 3385

55 step.1002d04.cons.ok

ATATTGATCTTTCAATCTACTGTTCAACATCTTTTATAGGTAATTGACTCT
 GTATATATGCTTTATGCTTTGTATAAGACGTTGGCATGTATGTTTGGATAA
 ACTTACCTGCATTCCTAAAACGTGGACGAGGAGAGCCGATAGGTTCTTTAT
 ACGTATCGTTAAATTAATCTCTATTTCCATAACTCACCTCAAATAATAA

TTCGTTAATTGTCACTCTGTTGTTGCAGTTCCTCTTTCTGAACAACCTTATG
TTTACGTTTAAAGTTTTCTAGTTCATCTTTCGTTACTGTTCCCTGAGAATGT
GTTTCTAAAGTGATGCCTGCATAGTTACCTAGTTTGAATGTATCTTCTCC
TAACGGCGTTACACTACACATCTCCCAACCGTCAATTTGATACAACATGTA
5 TTGCTTTTAAAGTCCGTCGATAAGCCCCATCTGGTTGCCTCCATTTTCGTTT
CATTCATGATTAATTCTGAACTTTTTTCATATTCGTCAAATGGTGATATCG
TTTTGTTTTCTAACAACGTTTAACTGCCCAGCTGACTCAATAAGCGTCT
TAGCTATTAAATGGGTGCTTTTGTATAATCTTCTCGATACATAACGCCTAACA
ACTTTTGATATTCAACTACTTTCATGTGAAGAACCTCTGCGTTTTCTTGTA
10 GTACTCAAACCTCAACTACACCTGTTTCTCCGTCCTTATTCTTTGCGATGTT
ACATTCAAACATTTGATCTGCTTGCCTGAGTCATCAACATCATCACGGTTGTAGTA
ATCATCTCGATATAACAACATAGCTAAACTTGCATCTGCTTCAATTCACC
TGCTTCTTTCATGTGAGATAGCATAGGTCTTTTGTCAATTTCTTGTCTTAC
ACCTCTGCTCAATTGAGATAGCAACACAATAATTGCACCTGTTTCATTGTC
15 AATAATCTTCAAATCTCGCGATATCTTTTCGATACCATTACGACGATCTAA
CTTACTGTCTGTCTGCATAAGTTGTAAGTAGTCAATGAAGATAACCTGTTG
CACATCTTTGTTCTTCATCGCTTGTTTACGTACATCATGTGTAGTAATATT
GCTTTTATCGTGTATATCTATATCAAGTTTGAGTATTCTGTCTGCTGCAGT
TGTTAAACGTGTTAATTCATCCGGTTCTAAATCTTAAATTTCTTTGATACG
20 AGTTAGTCTATCCAGTTTCTGCTGATAACATCCTTTTCAATACAGACAC
GCCAGTTGCTCTAGACTGAAGAATGAAGTTTTATAGCCTTGAGACGCTAT
ATTAAGCATCATATTAAGCGCAAACCCGTTTTTACCTACTGACGGTCTCGC
AGCGATTACAATCAACTGTGTAGGTTCTAAACCACCTATTTTGTAAATCCAC
CAGTTTATAACCTGTATTGATTTTTTGTGTTTGGTTCTTCGCTATATAATTC
25 TTCGACAAAGTGATCTACAAATTTTTTAGTCCCACTTCTTTCATTTGCACT
AATTAACCTGACCTTTTGTAGTTTGTAAAGCATTTTCATCAAAATTTTGTAT
ATTCGGATCAGAATTGAATCTTGTAAATACGTTTTCGCTACGCTCTATTG
ATAAAGATTGAGCAAATCTTGTGATATCTTTCAAAGAATCCGTAACCTAT
AAATTTTGAGTTATACAAATTTGAAATGGTGTCCATATCTAGGAATGACTT
30 ATCTTTTGTGGTTTTTAAATAGATTTCTGTTATGATCTACCTTACCGACTTC
GAATACATATTCATATAATGACTTCATACCATCATGTGAGAACATTTCCGG
TCTCACACGTAACCTTCTCAATTATGTCCGGTTTTTGAAGTAACTAGCAAC
GATTGTACTTTCGATTTTCATGACGTTTCATTCATCGTCACTCACTCCAACT
CACTTAACCTCTTTCTGAAGTCGTCTAATATCTTTTACGTTGTGCTACGT
35 ATTCTGGATCATTTTTTCATCTCCAACGATGTCTAGCAGTTTTTTCGTCGA
CAGGCTCTTCTTTTACGACTTTGACTTCTTTTCTCATTATGTTTGGAAATAC
TAGGTGGATAAGGATTAGCATCATGATATATTGCATTACTGTTTTTTTAG
TCGGTTCAATAATCCCAATTTTGGCTCAAAATGTTAACCCATGTTTCTAATT
TAGGCTGTCAAAGTCAATGTTGTATACATGTCTAATTGTCTTAATTACTT
40 CTAGGGCTTGTGTTTGTAGTCATAGGCATTAACTTTCATCTCCTAGCTCTTT
CTCCATTTGAGCAATTACATCATCTGTTACAGTTTTTTTATTTTTAGGTTT
AATTTTATTTTCAGCCTCTTTTTTATTCTTAACGTTCTCTTTAGCCAGTT
GTTTAAACCTTTGATTAGGTAACCTGCATGACAACCTTTGTCTTTTGTATA
ATCAGTAGCTACTTCAACAACCTTCATCTGCATGTTGTCCAATATCATCAAT
45 GGCATATCCTATCTGTTCCATTTGGTTAGGAGTTAAATATGAGTAAGGTT
ACTCATGATGTAATTAATTGAGTTTTTGAAGATATCTTTATCTACTTCTTT
TTCTCTTTCTTCTTACTTCTTTATCTTCTTCTATATCTGTTGCGTGACT
GTCACGTGACTTCACGTGACTATCTAAAAGTTTTGTTTTTTTCTTTGCTT
TTGTTTACGCAAACGGTTTTGTCTCTTATCTTTTCTAAACCTTCGATGTT
50 CTGATGTTTTTCCCAATTAGATACTTTAAAGACACCATTCACTTCTTCAAT
CATGCTTAGCTTTTCAATGTTTGTAAACGTAATCTTATTGAATTGATAGG
TCTATTAAATTCGTTAGCTAACATTTCTTTCGTTATAGGGTAGACTTTTCGGA
TAACATAATGTATCCTTGTTCGTTATACTTTCCAGCTAATGTGACGAACCTT
AACCATAAAGTGATGATTGTATCTCGTTCTGGCAGTGCTTCTATATACTT
55 GATTTTGTATCATCGAACATTCC

Sequence 3386
step.1002d06.cons.ok

AAACATCACCTTTAAATAACTGATTAATTTGATGAATAATATTAATTGAAC
TAAATCATCCGTTAAATCTCATGGAATATGCATAAGAATAATTCCAATTTT
TTAAATGTGCCAAAAAATAATCTAAAGATTCTACAGTTTGCGATATGTCAT
CAATAATATATCCATTATGCATATGTTTAAACATAATCTGTTCTCTTTTAT
5 TAATTTGTGGTATACCCGCGCCAATACAACAGATTTGTAAAAAAGTTTCG
GCTCTAAAGATAAATCTACTACAACCTCTTAATTTTGAAATAACGCTTACAA
GATCTTCTTCAAACGGTACTGTTTCAATATCAATGATTGTTTCTTTTAT
TCTGTATCACGCTCTCAATTACATCACTTATTTCTTTTCTCTTGGTGAT
ATAAATCATTGAGATGAGCAACTTCATCTATAAGATTTTCCGTAAAGATTAT
10 CTCTACTCTTAGTTAAATTTTCAACTTATAGCCATCTTTATGTTGAATAT
ATTGAAATAAGCTGTTTACAATCTCTCGTATTTCAATCTCGTCCAAACCAT
CTATCCATAATCCAATATAAGTGTACATAAATGACTACTTTGGTTAGGTA
ATGACTTCGCATCAAAAGGCGTTACCCCTTAAAGATTAAATATCATTATCATGT
TCAATCCTGCGTAGTGTTCAATCATATTTTGATTAGCTTGTGTATCAACGA
15 GACAATAATATGCTCTAGAAATAGAGTCATATAAATCTTGTGTCACCACTT
TATTTCTCTCAGTAAAAATAGAAAAACATAAAGACGATGCATCAATAGTGC
GTGCCATCATTTGAATTATGACGTTTCAATCAGATGCCACTATAACAGAATCAT
CTTCATTCAATTCTCTTTCTACATATGATTGGAATTTCTCTTCTATCAACT
CAGCCATAGATGAATACGTAACCCCTTTTAAATTTTGATTGGAAATTTTAT
20 TAATCGTTACTGTATTAACATTTAAGTCTTCAACAAATATTTCTTCCCAT
CTTTTGAAAAATAGTGTTTTATTTGTTATCACCATCAGGTGTATAAGTAC
GCACTGCCGATATAAATCCTCTGTATCGAATACAAATCGTCTTTGAATTG
TACTATACTTGTAACTCTTACCACATCAGGTACCCCTCTTGGCTGAAGT
GAATTTTAGAAAACGTGTTATCACCTGTAATCGCTTGAATTTGAAAGGGAG
25 TAAAAATAAATTCAGTGCCTTCTGGCCATGAAAGATCGCGATAATCAATAG
CTTGAGGCGTTTGATGTCCAACGCCCTGTATCTCATCAAAATACAGACCAAT
AATGACTTTCATACAAATCATATCGATGGAGGAATGTTCTAAGATATGGAC
TAAATTTTAAAACTATCAATTGATAATCCACATTATTCGAATATGCATTG
TCATTAACATAATCATATCGTCAAAATCTGTATACTGTTTTTTTAGATAGA
30 ATGGTCTTGAGGTACTTTCCCAACCATCTGTTACGGCTATACCAAGCTGGTA
TAAATCGTTTCATATTTCCAAACCTCCTCCCTACCAATATTTATTTAAAT
TTGCTTATAGCTGTCAAAGTATAAGTACGCTTTAATTACTATGGGAGGTTT
AAAGCACGGAACGTGTTCCAATATTTGAAAAAGGAACATCAAGCTATGTTGA
CCCTATCCCCAAGCATTGATTTTAAACAGCTATCGTTATCGCCTTTGCTAC
35 AACAGCTTTCTTTTAGTTCTTGCAATTTAGAACATATAAAGAAGTAGGCAC
TGATAACGTTGAGGTAAATGAAAGGAGCGCCAGAAGATGATAGAGAGTAACT
TGCTTGTTTTTAACTTTAGTCATACCTATCCTAACTGCAATTTTACTCATAT
TTATTTGGAAGCGCCCCAATAAATAAAGATATGTTGCACTTGTGCGGTACAT
TGCTAACATTTGGTCATTGCATTGATTAACCTTAAAGAATGTGCTACGAGATG
40 GACCAATCAAGGTTGAACTGGGCTCTTGGAAGCCCCATACAGCATTTGTAT
TTGTAGTCGATATTTTACGCTCTACTCATCATCACAAGTATTATTTATTA
CCATATTAGTCATTTTATACTCATATCGTTCTATCGGTTTAGATAGAGAAC
GTCATTATTATTATTTCTCTATCATGTTTATGTTGATTGGTATTATTGGCT
CATTTACAACAGGAGATATTTTCAACTTGTTTGTGTTCTTTGAAGTCTTTT
45 TAATGCTTTCATATTGTTTACTCGTTATTGGTACTACTAAAATACAATTAC
AAGAAACAATTAAGTATATTTTAGTCAATGTTGTTTCATCGTCTTTCTTTG
TCATGGGTGTGTCAGTTTTATTTAGTTGTAGGAACCTTTAAATCTCGCTC
ATATTAGTGAAAGATTGTCACAACCTTCTGTACATGACAGTGGCTTAGTCA
ATATTGTTTTTATTTTATTTATCTTTGTCTTTGCCACTAAAGCAGGCGTTT
50 TTCTATGTACGTATGGCTACCTGGTGCTTATTTATGCCCCCTCCAGTAGCGA
TCATCACGTTCTTTGGTGCACTATTGACTAAAGTGGGTGTATACGCAATTG
CGAGAATCTAAGTTTATTTCTTAAATAATACAGTAAGCTTTTCTCATTATG
TCATCCTTTTCTTAGCATTACTTACAATTATTTTGGATGTATAGGTGCGA
TAGCTTACTATGATACGAAGAAAATCATCCTTTACAATATTATGATTGCAG
55 TAGGTGTCATATTAGTTGGTATTGCTATGATGAACGAATCAGGCATGACTG
GTGCAATATATTACACACTACATGATATGTTAGTTAAAGCTTCATTGTTCT
TACTCATTTGGCGTCATGTACAAAATCACTAAAACGACTGACTTACGTCATT
TTGGTGGCTTGATAAAAGGGTATCCTATTCTAGGTTGGACATTCTTTATTG
CAGCGCTAAGCTTAGCGGGTATACCACCTTTTAGTGGTTTCTACGGTAAAT

TCTATATTGTTTCGAGCGACCTTTGAAAAAGGATTTTATCTAAGTGGTATCA
TTGTACTTTTATCAAGTTTAAATCGTGTTATATTCAGTCATACGTATTTTCT
TAAAAGGATTTTTCGGTGAAGTTGAAGGATATACTTTATCTAAAAAGGTAA
ATGTTAAATATCTAACAACTATCGCTGTTGCATCTACAGT

5

Sequence 3387

step.1002d07.cons.ok

10 CAAAATACTTTAATATTTTCAGCTGCAACATGATAATCTCTCAAATCCTCAT
CAAAACCAAGAGCAAGATTTGCGAGTAACTGTATCATAAACCTTTTCTATCA
ACTCATATGCGCGCAACTTATTAATTAACCCATTTCCCTCTACCTTCTTGAG
GTAAATAAATAATCATTCCACCATGTTTCGTCAATATATTTTCATTGACGCTT
CAAGTTGTGCCCCGCAATCACATCTTTGACTATGAAAAATATCCCAGTCA
15 GACAAGCAGAATGCATACGTACATTAGGATTGCTTTTAAATCTCCTTTAA
CAATAGCTACGATTTCTTCATCGCTATAATCCGTTGTAAATCCATACATAT
CAAAATGACCAAAATCAGTTGGCATCTTGACCTTTGCCTTAAGATTAACAT
TAAGTTCAACAGCCTTACGAAAAGCAACCAAACTTTTATAGTAATCATTT
TTAAATGGTGGCGTTCTTTAAATGACTGGAGATCTTCACCCCTTAGCCATTG
TCCCATCATCATTCATAATTTTCGCAGATTACTCCAGCTGGTTGTGCTCCTG
20 TTAACCGTGCCAAATCTACGGCAGCTTCAGTATGACCATTACGTGTTAACA
CACCATTCTCTTTTGCTATAAGTGGAAATAAGTGCCCCGACGATGAAAAAT
CTTCAGGATTAGTATTTTCATCTATGAGTGCTCTAGCCGTTTGTGTACGTT
CATGTGCACTGATTCTGTAGTAGTTTATAATGATCAATGCTTACAGTAA
AATGTGTGCCATAAAATATCAGTGTTATTTTGGCTCCATAGATTGTAGTTTAA
25 ATCTTTCAGCTATAGATTTTATCAATTGGTGCACAAATCAGACCACGACCCT
CTTTAGCCATAAAATTAATGGTATTATCATCCATCCATTCCGTAACAGCTA
CAAGATCTCCTTCATTTTCTCTATCTTCATCGTCAACTACAATAATGCTCT
CTCCATTTCTTAAAGCCTCTATAGCCAACTCAATTGTATCGAATTGCATGT
GATATTTCTCCTAAAAACCAATGCTCTAAGCTTTTCTTCTGTAACTGCG
30 ATTGATTTTGATTTCATGATGTTTCAACATATTTGAATAGTACGTCAGACT
CCAAGTGCACCTTTGTGCGCCCACTTTTATAGATGAAAGAATAGTAGATCGAC
GTGTTTCTGGTATAAGATGTATATCAAAAGTATAATCATGTAAATCAAATA
CAGTAAGACTAACTCCGCTTACAGTTATAGAACCCTTGCTTTACCATTTGAT
TCAAAATATTTTATGTTGTTTAAATAGAAATAATTTTGTAGTTAGCAGTTT
35 CATTGATTTTAGAAATTGTTCCAAGCTCATCAACATGACCTAACACGAAAT
GTCCACCAAACTCTCCACTTCCACTCATGGCTCTTTCGAGATTAACCTCTG
TATTACGTTGAACACTTCCAAGATATGTTTGTGTTTCAGTCCCTTTGATGA
CTTGAAC TGAAAACTTGAGTCAGTGAAATCTATCACAGTTAAACATGCAC
CGTTAACACTTATTGAATCACCAATATGCATATCAACTAAAATGTTTGTG
40 CTTTAATTTCAAGCGTTCTTACTGATTGTTTCAGAGCGAACTTGTTGTACAG
TACCTATTTCTTCAATGATACCTGTAAACATAGACATTCACCTCTTTCGTA
ATTTCAATTTTAAATTTTGATTAATTAACCTTGAATCAACAATTTCAAAT
GAGTTGCTTCAGGCAAAATCAATGACCTCGTCAGTCTTATAAAATGATGTT
TGCCAGAACCACCAATTAATTTTCGGGGCTATATATAAAATGAGTTCATTTA
45 GATGTTTGGATTGGAGAAATTGAGATGTAATATTTGGGCCTGCCCTCGACTA
GCAGTTTCCCAATACCTCTTTGATATAAGTCTTGTAATATTGTCGTTGTAT
CACAATTACTAATATTTATATTTTAAATAAACTTTTATTTGTTTAAAT
TTTCATTTTCAGTGTAATCCATATCTCCGATGCAGTATCTTTAAATATTT
GTTGATTAATAATCGAGTTGACCTTTCTTAGAAAGAATAACTCGAATCGGAT
50 GCTTTCATCAGGAACCTGGTTGTATACAATGGATTGCTGCTCAATGG
TTCTACGCCCAGTAATAACTGCATCATGCTCATGTCTTAATTGATAAACAT
CTTCTTTAACTTCTTTGTTGTATCCACTTACTTTTCAATTAAGTCTGTTG
CTTGTTTACCATCTAGACTAGATGAGACCTTTACAGTTACTTCTGGAACCT
CGTTTCTTTTAGCAGTAAAAAGTCACGGTATAATGCAGCTGCATTTTCAT
55 TATATTGAAATTCACCTCTATACCAGCTTCTCTCAGAATCTCGTCACCCCT
TACTTACTAAAGTAGTATCTTTAACAGCATAGATGACCTTAGATATGCCCCG
CTTCAATGATTTTATGCACACAAGGTGGTGTGAACCATGGTGTGTGCAAG
GTTCTAATGAACGTATATGGTAGCACCTTGGGTATTTAAACCTGCCATTT
CAATAGCTTGTACTTCGGCATGTTTATCTCCCTTTTAAATGTGCACCTA

AACCTACAATCCTACCGTTTTTAACAACAACGGATCCTACTGGTGGATTAA
 CACCTGTTTGTCCATTTACCATTTTGTAGTTGAATAGCATCATCCATAA
 ATCTACTCAATTGATCACCTCAAAAAAACCTTATACCTATATTGTAGGCA
 TAAGGAGATATCTATATCTTTAAATAAGCCTATTAGCATACTTAAAAATGC
 5 ACATAAGTATACAATACGCTTATCATAACTACACGAATTTGTGTAAGTCGA
 TAATTC AACTTCTTCTCCCATCCAGACTCTAAGTGTGGCTCTAGATTCA
 CACTAGATCAGCCACTAACATCAAAAAACAATATTAGTGGGTGCGAGGCTCA
 ATTTACTGCCGTTGGGAATTTACCCCTGCCCCGAAAGAAATATATGAAAT
 TGTATGATTAAATTGAGTCCAGTAGTTCTTCACTACATTTGAAATCTTACT
 10 ATATTTAAATTTAAAAATACAATTTTATAGCTTACTTTATTTAAAGATTTTA
 TGAGAATATATTAAGTCAACAAACAGCTTAATGCATATTACTTATTTCTA
 TTTTCTGTGTTATATATAATTGAGATACTGTATCAACAACCTTGATTAAAC
 ATCATCTTTACGCCATTTCTAAGTTTATTGACACCGTTGGTCATTTGTCCA
 ATTGCAATCACATTTTTTAAAGTTCCATATCTGGGGCTTATGACTTGATTT
 15 GTTCTCGGATAATTTGAATGCCACCCATAGGATGACGTTGAACAATTTGT
 CTATTTTCTAAGTTTAAATTAATTGATCATCCTCATCTAATTGAGAAAGA
 TGTGTTTATAGAGCCCGTTGCATTAATAACGATGTCAAACAACCTATAATTT
 TCGCTAGTGTGCTTATACTTCAACATTAAGAGGAATCCTGCACAAAAATGT
 CTCCATTCCGGTAGCCATTTTGTGTTAAAAACATTGATTGACAGTGTCAAAT
 20 AGAATTCCTTTTATGCGCTGAAATTTGTTATATGGAGACTGAATGTA

Sequence 3388

step.1002d08.cons.ok

25 GATTTAGAACAGAATGTTACAGTTACACAAGTTAATGATGCATTTAAAAAT
 GCCGATTTTATCAGGTGTTCTTGATGTTGAAGAAGCTCCTTTAGTTTCTGTA
 GACTTTAACACAAATCCTCATTCAGCAATTATTGATTCTCAATCTACGATG
 GTTATGGGACAAAAAAGGTGAAAGTTATCGCTTGGTATGATAATGAATGG
 GGTATTTCGAATAGAGTTGTTGAAGTAGCTGACAAAAATGGACAATTAATT
 30 GATGATAAAGCAATGGTAAAAGCCATTTAATTAAATATTTATACTTACACT
 AAACCTGAGATAGCCATATTATCTCAGGTTTATTTATATACAAATAGCCCT
 AAACAAAGTTGAAGCACCATTGATGTTATTATAGATAACCAATTAGCTTTA
 GACATATTTCTAGATACTTTCTTTTAAAAAAATGATCCATAGTCATTATT
 TATTAATCAAGTTGAGATGATTTTGAAATGAGGTGATTTTAACGAATCATC
 35 TTTCAAAATATAATTTAATAATGGTACTTTTGGTCGTATTTATAAAATGTT
 GTAGTACATTTAAATGGTTGAAATTCATTTAATATGGTTAATCACGTTAAA
 ATGTATTATTTTAAATAAATACATAGCATTCTAAAACCTTTATTTATAAG
 ACAATTTTCCTTTTTTACACATAGTGGTATTAACAATGTATCATTTACATT
 CTTTCGTTTATAGTTGCAATTGACCTTTCTTTTTTACACAGTACTTTTATA
 40 GTATAATGTACGGTGATTAAACTTGGGAGGGTGACATATATGAAATGCCCA
 AAATGTAATTCTACACATTCCAGAGTGGTTGATTCAAGACATGCAGATGAG
 GCCAATGCGATTAGACGTAGAAGAGAATGTGAAAATGCGGAACGCGTTT
 ACAACATTTGAACATATTGAAGTTAGTCCATTAATAGTAGTGAAGAAAGAT
 GGGACTAGAGAACAAATTTTAAAGAGAAAAAATATTAATGGTTTAGTAAGA
 45 TCTTGCGAGAAACGACCAGTACGTTATCAACAACCTGAAGACATAACTAAT
 AAAGTGGAGTGGCAACTTAGAGATGAGGGACAACTGAAATTTCTTCTAGA
 GAAATTTGGAGAGCATGTTATGAATTTGTTAATGCATGTTGACCAAGTTCC
 TATGTAAGATTTGCATCTGTATATAAAGAATTCAAAGATGTTGATCAACTC
 TTAGAGTCAATGCAAGGTATCTTGAGTGATAATAACGGAGTGATAAATAG
 50 ATGGGGTTACAAACCTATGAATATGGTCTAAAACCACAAGATGGATTGAG
 GTGATTACACATTTTCGAATTCACCTCACAACTTTAGATATTTTAAATCGA
 CTATTCACCCCTTTAATCGGAGTTGAATCAATTGGACTCTATCATTTTATG
 AGTCAATTCATAGATAAAAGTCAACAACCTCGGGTTAACGCATTATATATTC
 ATGAATGAACFAAAAAATTAACCTTATTAGATTTACAGGAGCAAAATGGACAAT
 55 TTAGAGGCTATTGGATTGATTAAAAACATTTGTAAGGCATGAAGAAAACCTAC
 TCTCACTTTGTTTATGAGTTAATTCAGCCTCCAACAGCCTATCAATTTT
 AATGATCCTATGTTATCAGTATTTTATTTAGTGAGGTTGATAAAAAACGT
 TATCAAGCACTTAAATCTTATTTGAAAAAGATGAGAAAGATTTAAGCAAA
 TATCAACAGACAACTAGAAAAATTTACAGAAGTATTCACGTACCTAAAAAG

GTCAATGTTTCTGATCAAATTAATTTAAAGCAAATCAAACACTATGATGGT
 ATAGATTTATCTAATGAACTTTTGATTTTGAAATGTTGAGACAGATGTTG
 AACCATCATTTTATTAGTAATGAAATTATCGATAAAGAAGCTAAGAATTTG
 ATTATACAACTTGCGACACTTTATGGAATTACTGAAGATGGTATGAAAAAT
 5 GTTATATTAAAGTTCCATTACCAGTGCACAACAATTATCTTTTGAAGAAATG
 CGTAAGAAAGCTAGAACTTATTACCTGATTGAACATGATAATCAATTACCA
 AAATTAGAGCATCAAACAAAATAAAATTAACGATGAAAAAAAAGATCGACAA
 GCGGAAGATACAACAAATGATTGGTTACAACCTGCTAGATGAAACAAGTCCG
 ATTGATATGTTAGCAAGTTGGTCTGATTGCGAACCTACACAGTCGCAAAAA
 10 AGTATGATAGAAGAATTGATTAAACCGTGAAAAAATGAATTTTGGTGTAATC
 AATATACCTTTTACAGTTTGTATTGTTAAAAAGAAGATATGAAGTTGCCAAAA
 TCTTATATTTTGAATTTGCTTCCAACCTGGAAGAAAAATGGTATTTCAAAT
 GCCAAACAAGCATATGAATATGCATTACAAGTTAATCAACCTAAAAATTAC
 GAAACACATCTTAATGATAAACGACAGAACAATCGTGGAGACAAAATCAA
 15 TTTTATCCAAAGAAAAGACACCTAAATGGCTTCAAAAATAGGGACGATCAA
 GAAGAAAATAAGAAAATAATGATGACACTCTCGAAGAAGATCGACAAGCA
 TTTCTTGAAAAGTTAAATCAAAAGTGGAAGGAGGAAGATAAATAATGAAAA
 GTTTCAGAAATATCATGGGCGATTCTCAAAATCTAGATAAACGATACAAAA
 AAATAAAACAAAATGTAATCAATGATACTGACGTTAAACATTTTCTTGAGA
 20 AAAATCGTAGTAATATAACTAATGAGATGATAGACGAAGATTTAAATGTTT
 TTCAAGAGTATAAAGATCAACAAAAAGTTTATGATGGACATCGCTATGATG
 ATGTGCCGAATTTTGTAAAAGGACATGTTCTTGAACATATATTGAAAAATG
 AAAGAATCAAAATTAGATATCTACCTTGCCCGTGTAATAATTAACATGATG
 AGGAACGATTTGATTACAACTTATTACATCTCACCATATGCAAAGAGATA
 25 CACTTCATGCAAAGCTCAAAGATATTTATATGAATAATCGAGAGAGACTTG
 ATGTAGCAATGGCAGCTGATCAAACTGTACAGCAATTACTAACGATGAAA
 AAGTAAAGGGGTATATTTTATATGGTCCTTTTGGTACAGGAAAATCAATCA
 TATTGGGTGCTATTGCAAATCAACTTAAATCGCAAAAGATTTTCATCAACAA
 TTGTATATTTACCAGAATTTATTCGCACTTTAAAGGTGGCTTTAAAGACG
 30 GTAGTTTGTAGAAAAAATTACAACGTGTGCGAGAAGCTAATATTTTGATGT
 TAGATGATATTTGGCGCAGAGAAGTCAACACCGTGGGTAAAGAGATGAAGTGA
 TTGGTCCTTTATTACA

35 Sequence 3389
 step.1002d09.cons.ok
 CAATGGCTTTGGGAAACTTCACTTTATCAAAAATGGTTGATTTCTTTTATT
 TTAATTTGCAATATATTTGGTACGATATACGGTTATCATTGGTATAGTAGT
 CAATTAGCTAACACACCAAATTATTTTAAATTATTTGTTCCAGATAGTCCC
 40 ACTGCTACATTATTTTTATGTTTATCATTATTATTCATACTATTTAATAAG
 CGTTATGCCCTGATTGACGCTTTAGCTTTTATAACTCTATTTAAATATGGT
 AGTTGGGCTGTAATTATGAATATTTTGATGTTTATAAAAATGAATGATATT
 GCCATAAACGGACTTATGTTGCTCATTTCTCACGCAATTATGATATTAGAA
 GCTATCTATTTTATCCTCGTTTFAAAATATCTAAATTTGGCTGGATTGATG
 45 AGTTTCATATGGGTGACGATCAATGACGTAATAGATTACATATACGGACAA
 TATCCCTACTATGATTTTATCGCCAAACATTTAATTGAAGTAGGGGTATTG
 GCTTATAGTCTCACTATCATTTCGTATATTTTATTTTAAATTAACAAAAG
 TGGTTGAAAGTTAAACATTTGATTAAAGACAAATTTGACAGTAAATGAA
 AATAAAGGAGAGTGAGTTTAAATGCTTTAATATCAATGATTATATATTTG
 50 TAATATTAATGGTCAATCCAAATGTGGGCTCAACACAAAGTTAAATCTAATT
 ATGAAAAATACTCACAAGTGAGATCCACAAGCGGTAAAACTGGTCTGTAAG
 TTGCTGAAGAAATTTTACATGCTAATGGCATTACGATGTTGACGTCGTTA
 AAGGCGAAGGTTTCTTAACTGATCACTATGACCCAAATAAAAAAGTTGTTT
 GTTTATCTCTGCAAACTACGATAGACCATCAGTTGCTGGTACTGCAATTG
 55 CAGCCCATGAAGTAGGACATGCAATTCAACATCAGCAAGGATACGCACCAT
 TAAGATTTAGAACAGCATTAGTACCATTAGCTAATATTGGTAGTTCTTTAA
 GCTACATTATTATTATGGTAGGTATTGTTTTAACAGCTCTCGGTAGTGTCT
 TTGGCTCTACAGTTCTATGGGTTGGTGCTGGTTAATGTCACCTTGCAAGTTT
 TATTCTCTATCGTGACATTACCTGTTGAGTTTAAACGCTAGTTCAAGAGCGA

TGAAACAAATAACAGCTTTAAATATCGTTAATGAGAAAGAATACAAACACG
 CGAAAAAAGTTTATCTGCAGCAGCAATGACATATGTGGCATCTACTGCAG
 TTGCGTTAGCTGAACCTTGCTCGTATTATATTAATTGCAAGATCAAGTGATT
 AATCATATATAACAACGCAGTTCCCTGGTAATCATTACAATGGATTTGTAAT
 5 GGTTACCTTTTTTTATTTGAATACGTGACAAAATCCTATCTTTTAATCCAT
 ATTAACATTGCACCTTGAAATGCAAGTCACCTTTATTTGTAATATACTATA
 TAAGAATAGCGATAAAGGAGTACTTCTATATGAAATCTATGGAGCAAAATGC
 AACAAGAAGTTGATGACTATATCAGTCAGTTTAAAGCTGGTTATTTTCTC
 CACTTGCTTAATCTAGCAAGAATGACAGAAGAAGTTGGAGAATTAGCACGCG
 10 AAATTAACCATCATTTTGGAGAGAAAAAAGAAAGATACAGAAGAAGATA
 AATCAATTAAAGCTGAATTGGGAGACAATTTATTTGTTTACTGTGTATTG
 CAAACTCTCTAAACATAGATATGACAGAAGCTTTAATGACACGATGAATA
 AGTTTAATACAAGAGATAAAAAATCGCTTTGAAAGAAAATAATTTAAGAAAG
 GATGGGCACAAATGAAATCGGTATAACGTGTTACCCGTCCATGGGTGGTT
 15 CTGGCATTTATGCTACAGAATCGGTATTAATAATGGCAGAACGTGGCCATG
 AAGTTCACCTTTATACCTCAAACATACCCTTTAGAATACGCAAACCTTTAC
 CTAACATGACGTTCCACCAAGTTGAAGTCAATCAATATGCCGTATTCCAAT
 ATCCACCATACGATATTACATTAAAGTACAAAAATTTCTGACGTTATACAAG
 AATATGATTTAGACATATTGCATATGCATTATGCTGTACCTCACGCTGTAT
 20 GTGGTATATTAGCGAAACAAATGTCAGGTAAAAACGTCAAAATTATGACAA
 CACTACATGCACTGATATAACTGTGTTAGGTTATGACCATACTTTACAAA
 ACGCGATAAAATTTGGCATAGAACAAGTGATATTGTAACAAGTGTTAGCC
 ATTCTCTAGCACAGCAAACCTTATGAAATTATCAATACTAAAAAGGAAATCA
 TCCCTATATATAATTTTATAAGGGAAAATGAATTCCCAACTCGGCATAATG
 25 AAGAATTTAAAGATTGTTATGGTATTTACCTGAAGAAAAGGTATTGATAC
 ATGTTTCTAATTTTCAGAAAAGTAAAACGTATTGATACAGTGATTGAGACAT
 TTGCAAAAGTTTCATGAGAGTATACCATCCAAGTTGATACTTTTAGGAGATG
 GTCCAGAAATTAATCGATATGCGACATAAAGCAGCAGAACTTGATGTTGAAA
 CACACGTACTCTTTTTAGGCAACAAAAATGACGTAAGCGCATTTCTACCAAC
 30 TATCTGATTTAGTACTACTCTTAAGTGAGAAAGAAAGTTTGGATTAACTC
 TCTTAGAAGCAATGAAAACAGGCGTCTTACCTATAGGGAGTCATGCAGGTG
 GTATTAAAGAGGTCATCAGACATGAAGAACTGGATTATAGTAGATATAG
 GGGATAGTACACAAGCTGCAAAATATGCTATTAACTTTTATCAAATCCAG
 AGTTATATCAAAAAATGCAATCACAATGCTGAAAGATATTGAAGCAAGAT
 35 TTAGTTTCAGATTTAATTACTGACCAATATGAAAATATTATCGAAAGATGC
 TAGAACAAGGTGAGAACAACAATGAGTCATGAGATATTTGAACAAGCAAAA
 CCTATTTTAAAAAAGATTCAAAATAAAGGTTTTAAGGCTTATTT

40 Sequence 3390

step.1002d10.cons.ok

TTAAAGATTTGAAAGAAGTTTCTGACGCAGATATCAAAAGTTTGAAGAAC
 GTTTAGATAAAGTCGATAATCAATCAAGTATTGACCGTATTATAAATGATG
 CAAAAGATAAAAAATATCATTTTAAATCGACAGACTCTAGTGCCACATCAT
 45 CAAAACTGAAGATGACGATACATCTGAAAAAGATAATGATGATATGACTA
 AAGACTTAGATAAAATACTGTCTGGATTTAGATTCAATTGCTAAAAATGTTG
 ATAACCGTCAACAAGGTGAAAAATAGTGCTTCTAAACCTAGTGACTCAACAA
 CCGATGAAAAAGATGATTCAAATAATAAAGTACACGATACAAATGCTAGTA
 CACGCAATGCAACTACTGATGATTCTGAAGAGTCGGTTATTGATAAATTAG
 50 ATAAATCCAACAAGATTTTAAAGTCTGACTCTAATAATAAGCTTTCTGAAC
 AAAGCGATCAGCAAGCATCACCATCTAATAAAAAAGAAAATAACAAAGAAG
 AATCTAGTACGACAACAAATCAATCCGATAGTGATAGTAAAGACGATAAAA
 GTAATGATGGTTCGTCGCTCAACATTAGAACGTATAGCATCAGATACTGATC
 AAATTAGGGATTCAAAGATCAACATGTCACAGATGAAAAACAAGATATAC
 55 AAGCAATTACACGTTCACTACAAGGTAGTGATAAGATTGAAAAAGCACTTG
 CTAAGGTACAATCTGACAATCAATCACTAGATTCTAATTATATAAATAATA
 AATTAATGAATTTAAGATCACTAGATACAAAAGTAGAGGATAATAACACTT
 TATCTGATGATAAGAAACAAGCGCTTAAACAAGAAATTGATAAGACTAAGC
 AAAGTATTGACCGACAAAGAAATATTATTATAGATCAACTCAATGGTGCTA

5 GTAATAAAAAACAAGCAACCGAAGATATCTTAAATAGTGTTCAGCAAAA
 ATGAAGTAGAAGACATAATGAAACGTATTTAAACAAATGGCCGAAGTAATG
 AAGATATTGCTAATCAAAATGCCAAGCAAATTGATGGTCTTGCAATTAACCT
 CTAGTGATGATATTTTAAATCAATGTTAGATCAATCTAAAGATAAAGAAA
 GTTTAATTAAACAATTGTTGACGACACGACTTGGTAATGATGAAGCAGATC
 GTATTGCTAAAAAATTGTTAAGCCAAAACCTTGTCGAATTCTCAAAATCGTAG
 AACAATTAAAACGTCATTTCAATAGTCAAGGAACAGCTACAGCTGATGATA
 TATTGAATGGTGTGATTATGATGCTAAAGACAAAAGACAAGCGATTGAAA
 CAATATTACAAACCCGTATCAATAAAGACAAAGCTAAAATTATCGCTGATG
 10 TTATTGCGCGTGTACAAAAGGACAAATCAGATATCATGGATCTCATTCACCT
 CTGCGATTGAAGGCAAGGCAAATGATTTATTAGATATAGAAAAACGAGCAA
 AACAAGCTAAGAAAGATTTAGAATATATTTTAGATCCTATAAAGAATAGAC
 CATCCTTGTTAGATCGTATTAACAAAGGTGTCGGTGATTCTAATTCAATAT
 TTGATAGACCAAGTTTACTTGATAAACTTCACCTCAAGAGGATCTATTCTTG
 15 ATAAATTAGATCATTCCGCCACCGGAGAATGGATTATCTTTAGATAATAAAG
 GTGGCCTTTTAAAGTGATCTATTTGACGACGATGGTAATATCTCATTACCAG
 CGACAGGTGAAGTCATAAACAACATTGGATACCAAGTGGCTGTTGTACTCA
 TGTCATTAGGTGGAGCGCTCATCTTTATGGCGCGTAGAAAAAACACCAAAA
 ATTAATTTAAAAATTATATATAGTAATAGGCATCATCAATCTTATGTGTAG
 20 TTAAACTATTTTACAGTTAGTGATTGAGGTGGTGCAACGCTTACTCAAAGT
 TTATAACAAGAGTCAGTATAAAGAGTAGGACTAAGAATTTAAATCAATT
 CTGTCCTACTCTTTTTTGACATGCAATAATGATTGTTATATATAATGAATA
 ATATACTTTTAGAAATCACTTGAGTAAATTATAAAAAATGGACTTTTCAATG
 TAAAGTTACTCTAATATAGTTATTGAATATAAATAAGCACAAAGATGTATT
 25 ATAAATTTGAATACATCTTTGTGCTATTATTTTGTGGATGGTACCCTCAAAT
 AATTTTAAATTTTACTATTTACCTTTAATGTAATCTAAATCTTGAGCATC
 TTCTGCATCAATAGTTCTGAATGAAATGACACCTAATCCTTTAACATTTGA
 TTCAGAAATAACAATTGAACCGTCGTCATTAACTTTTCAACGAAAGCAAC
 ATGACCATACTGTGTATCAGCACCTAATTTGCCAGCTTCAAACACAACCTGC
 30 AGTATGATTTTTAGGTGTGTGCGTTACCGTATAGCCTTCACTTTTACGACG
 ATTATTCCAATTATGAGCATCACCTAAGTCAACAGAAATGGATGCATCAAA
 TTGATTCATACGGTGGTACACATACCAAGTACATTGGCCATGTGGGTAAGG
 TGATGTACCGGAAGTTTCAAGTAAATGGTTTGAAGTCATTACCAGAGACATC
 CGTACCTATTGATTTGTGTATTTCTTCATGTTAGGCATTTTTTCTTTGTC
 35 AAAAGATGTTAAATGATAATGTTTAAATAATACTATTTAATTTTTTAGAATA
 ATTAGGATCTGTGGCGTATGAGCGTGACAGATGTGAAGTAGCATCTTTATA
 TGATAGAGCTTCACTTTCCAAGTTGGTTTATAAATTGACGGATTACCATC
 GATACCATGTTTTGATTAAATCTGCATAATCTTCAAGAGATTGTTTAGTACT
 TGGGTATTTACGGAACCTGCTTGGATACTGAACATATGATTACTGCTATC
 40 AGCTTCTAAAGTATTAAAAGTTACAGATTGTCTTTGTAGTCACCTTTGAT
 TCCAAACAAGTTATGATTGGTGATTGTGCAAGTGAACCTTTTCCAGAGTC
 AGATTCTAAAATAGCTTGAGCAATCATAACTGATGCATATATATCTTGGTC
 TTTTCCAATCTGATGCGCATCTTTAGCAATTGATTGAATGAAATCACGTGT
 GTCTTGACTGTCAACAACATTAAATGAATCAGAAGATAAGTCAACCATGTGTC
 45 TAATTCAGGTAGTTGTTGGAAGAAAGACTTGTGTAACGTGATTGTGAGCGTTT
 AATGTCATCCTCAAATGATTGTGCGAGGTTTGGATTGATGTTTTAATTCATC
 ATCTGTTGGTAATTGTGGATTCTTGTGCGCATTATCACTTTGAGATGTTTT
 AGTGTGACTTGTATTCTTGCTCTTATTGTAATCTTTTTCTGTTTCTTAGC
 GTCTGACTATATTCATCTAATATAGAGTCTAATGCTGAATCTGTAATACT
 50 AGAATATTTCTTGTTACGTTCTGAGCTATCATCATTCGCTTGAGTTGTTG
 TTGCTCAGAAGAAGATTGATTTTTGTTTGTATCT

Sequence 3391

55 step.1002d11.cons.ok

TCATAGGGCTTAGATAATGCTTTATCCAAACCCACCTTTCCATCATCTCT
 ATATCAACAGTATAATTTTCATGAGTTAACTCAAGCTCTATAAATCTAGCA
 AGATTTTGCTCATCTTCTACAATTAAAAATATTGTGCATATTTGCACCTCAC
 GCTACATCTTAAACAATTAACTCGTATTTCTAAAGATGTAATTCAAAATTG

TAAATGATTTGTTATGAACTAGTATCATTTTTACCATGAACCACGATAATA
GATAGTCGAATAAGTAGTCTGTATTTAGAAAAAGTCACTCTAGTATTATCA
TTGTATTTAAGTTAAATGGAAAAAATATATATTTCTAATTGACAATCATT
ATCAATCATGTATAATGATAATTGAAAATGATTATCAATACCAATTGAAAA
5 ACATTCCCCCACATACAAGTTGTTCTTTTGGATTGGTCATTTTCAACTAT
CCCCTTTATATGCCCCGTAAAAGACTAACGTTAAGAAATGACGTTTCAATA
AAAGCAGTAGACCTTTGACACTTGAGGTCTGCTGTTTTTCAATGATAATA
TAATGAGAGAAAAGTAAATAAATAAGTGACTAATAATCTTATCAAAATTTTT
AGAACGCTCATCACTTTTGATTAATGAAATAATATATACAAAAGAGTCTGT
10 AACACATGAAGTGCACAGACTCTTTCCTACGATTAGTCAAGTACTTATAAT
TAAAATAAGCATTACTTAACATAACCGTTCATTTTAAATTTTTTGAATAAG
GTGTTAACTATATTTGATTGTACGTTTGGCTAATGTGGTGACATTTGATA
ATCGTACAATTTATGATTCAATACAAGATAATGAATTGTTTTTAATAATCG
ATTTATACAAGCAATGATGGCAGTCTTATGAGGTTTCTCATTAGGCTGCTT
15 TCTTAGTTTGTAGTAATAATCGACGACATGATTGTCATAATGATGCTGCCC
TCTTATTATATTCATAATCACCCTAAATAAGTTTTCTCGCTTTTTTATT
ACCACGCTTGTGATGGTATCTCTACAGTGTGTATGACCTGATTGATATCG
TTTGATATCAATGCCAACAAAAGCATTGAGTTGTTTATTTGATTAAATCG
CTTAATATCACCCTCTCCCCAATAATCATAGCTGTGCTTAGCTTACCAAT
20 ACCAGGTATCGAATGAATATTTTCAAAATAATCGAGTTGTTGTGCTAATTG
AATCATGGCATCATCTAATTGTTTGAGATGATGAATAGATTGTTTAAATTG
TTGAATAAGTAAGCGTAATTTTTCGACTAGAAAGGAATGCTATCGACATT
AGGATAGCTTTCTTGAGCAATCACCTTAATTGAAGTGCATATTTGTAGC
TTTATCCATTGACATTTCCCTTATCTGTAGAATTGAATATATGTGTAATCAG
25 TACCTCCTTGTCTGATATCAAGAACCATGTCTGGATGAGTAAAGATTCTGTC
GATGTTGAGTGCAATGATTGAATATCGACTACTAAATAATCTTTCTAAACC
AGGGAATGTTTGATGGAGTAATCAAGGATCTGAAATTTAAGTCGATTTTG
TTCATTTCTGATTTCTAGATGAAAACGGACGCTTCTCTTAATTCAAAGAA
TATTAACTCATGTATAGGTAAGTTGTCTGTTTGTTTAAGCGTCGGTCTTAA
30 ACAAGCAAGCTTATGAGCATCTGCCTGATCAGTTTCCATGATCTTAGAGC
GCTCGTTTTTAAATTTGGCTTCTAACGGATTCAATTGAATATAGTTAATTTG
ATTTACACAACAAAATCGTTCCATACTTCTTGAATAGATACCTGTAGATTCT
AAAAATGAGTTGTGGGTGGTCTAAGTCATTCAAATACTTGAGTAAATAGTT
GTAACCATTTTATTATTTTGGATGAAAACTCTTTTGGAAATTTTTCATT
35 TTTATAATGTGCAACTACACTACTTCTTTTACTAATATCAACACCTAAGTA
ATCGATAAAAAAACCTCTTTGAATAATTGAGAAGCTAAAAACTTTACTTA
ACCTTTCTCATTTTCAATTTTCTTATACACGGTTTCAAGAACCACATACTA
CAAAAGCAATTTCAAAAGGCGAGAGTAAAGCTGACTTTGTTTTTATACGGAT
TTAAAATCCAAGAGTCTGGACAGTCTACTTCTCTCTATAACTATAAAAAAT
40 AGCTATGAAAAATCTATCGTCATAGATTTCTTCATAGCTAATCTTAGTAT
GTTTTTATTTTCATAATTAATTGTTATGTTGCTAATTTTGGATGGTTTTCA
ATCTAATCAGCTTCGTTAGTCACATAATATGAATGAACCTAGGTTCAAACCT
AAAAGTAGTTGAGTGTTATCAAAAGAAAAAGAAAGACCCTACTTATAATG
TAAATATCAAAGGGACTGAAAAAGCTAAAGATGAGAAATTTAAGTTTATTA
45 CGATAAAAGACTATTTAACTTGCTAGATTATTTCAAGAAAAGAGATGAAG
AAAGTTATGTTTTGCTATATCTATTAGGCATTACTGGCGCAAGATATAGCG
ATGTCATTTATATGACTTACAAAGATCTAAACAAAGCGAATGGCATAATTC
ATTTGCCGTGGAACGAAAACAAAGAATTCAAAACGTGATGTAGAAGTTAATT
CAAAAGATATCATGCACATAAATTCAAAATTAGCTAAAATGCCGCGTAGAA
50 TTGATGGCAAGTTATTTCTCGGTTAGTCATACATCAGTAAGTAAAGCGTTTA
GAAAAGCAAAAGAGTATAGGATTAAACGATAATAATATAACTCCCTATT
CACTCAGACATACGCACACATCTTACTTACTATCTAAAGGCATACCAATCG
AGTATATAAGTAAACGTTTAGGTCACGCTACTATATCACAACGTTAGACA
CGTATTCACATTTATTAGAAGACATAAAAAAGAGCAAGGCCAACGTGTCA
55 GAGAAATATCTCTTTGACACTTATTTGACACTTACTCTCTCGAAAGCCCGT
CATATCAACGGTATAGTACGGAGAGTTTCATTACTTCTTAATGGTGGATTCT
ATTTTGTAGTTTGAATCATCGCCAGGAACATCATTTTTCAGTTAATAATAAA
TGATGTGGTGTAACTTCTGCTGTACATGGATACCAGCTTTTTTAGCGTCT
CTGAT

Sequence 3392

step.1002d12.cons.ok

5 TCAAGATTGGAAGTTACGTTCATGTTCTAATTGAAATGCATGGTAATAATGG
TAGTATCGATAACGATCCTCCAGCTGCTATGCGTTACACAGAAGCTAAACT
TAGTCAATTATCAGAAGAACTATTAAGGGATATTAATAAGGAAACAGTATC
ATTTATTCCTAATTTATTAATTAATGGATCTACGGGGATTTCTTCAGGATA
10 TGCTACTGATATCCCGCCGCATAACCTCGCCGAAGTAATACAAGGCACATT
GAAGTATATCGATCAACCTGATATTACAATTAATCAACTGATGAAATATAT
CAAAGGGCTGACTTTTCTACAGGTGGTATCATTCAAGGAATAGAAGGTAT
AAAAAAGCGTATGAGACCGGTAAAGGAAAGGTTGTCGTGCGTTCACGAGT
AGATGAAGAGCCTTTAAGAAGTGGACGTAAACAATTAATTGTGACTGAAAT
15 TCCGTATGAAGTGAATAAAAGTAGTTTAGTTAAAAGAATTGACGAATTACG
TGCCGATAAAAAGGTTGATGGTATTGTAGAAGTTTCGAGATGAGACTGATAG
AACTGGATTACGAATTGCAATCGAATTAAAAAAGATGCTAATAGCGAATC
AATCAAAACTATTTATATAAGAATTTCGGATTTACAAATTTTCATATAATTT
TAATATGTTGCTATTAGTGAAGGTGCGCCCTAAGTTGATGGGATTACGTGA
20 AATTATAGAAAGTTATTTAAATCATCAAATTGAAGTGGTTACAAATAGAAC
GCGTTATGACTTAGAGCAAGCTGAAAAACGTATGCATATTGTGGAAGGATT
AATGAAAGCTTTATCTATACTTGATGAAGTTATTGCATTGATACGTAATTC
TAAAAATAAAAAAGATGCTAAAGATAATTTAGTTGCAGAGTATGACTTTAC
TGAAGCTCAAGCAGAAGCTATTGTTCATGTTACAGCTGTATAGATTAAACAAA
25 TACTGACATTGAAGCTTTGAAAAAGAACATGAAGAGTTAGAAGCTTTAAT
AAAAGAAATTAAGAAATATCTTAGATAATCATGAGGCACCTTTTAGCAGTAAT
TAAAGATGAACATAAATGAAATTA AAAAGAAATTTAAAGTGGATCGACTATC
TACAATCGAAGCTGAAATTTCCGAAATCAAAATTGATAAAGAAGTTATGGT
GCCTAGTGAAGAAGTGATTTTAAGTTTGACGCAACATGGCTATATAAAACG
30 TACATCTACACGTAGTTTAAACGCAAGTGGTGTGACTGAAATCGGTTTGAA
GGACGGCGACCGTTTATTA AAACATGAAAGCGTGAATACTCAAGATACTGT
TCTTGTATTTACAATAAAGGTAGATATTTGTTTATACCTGTTTCAATAAAT
AGCCGATATCCGTTGGAAGAGCTTGGTCAACACATATCACAATTTGTGCC
AATAGATGAAGATGAAGAAGTGGTAAATGTATACAACGAAAAAGATTTTAA
35 AAATGAAGCCTTTTATATTTATGGCTACAAAAACGGCATGATTAAAGAAAAG
TAGTGTCTTACAATTTAAACTACTCGGTTTAAATAAACCACTCATAAATAT
GAAGGTTAAAGACAAAGATGAACCTATTATATGTCGTTTCGATTAGAGTCTGA
TCAGTTAATTACTGTTCTAACCATAAAGGCATGTCATTAACCTTATTCAAC
TAATGAATTATCGGATACAGGCTTAAGAGCAGCTGGTGTAAATCAATTAA
40 TCTTAAAGATGAAGACTATGTTGTTATGACAGAAGATGTGAACGACTCAGA
TTCCATAATAATGGTTACACAACGTGGTGTATGAAGCGTATTGATTTTAA
TGTTCTTCAAGAAGCTAAACGCGCACAACTGGAATTACTTTACTAAAAGA
ATTAAGAAAAAACCGCATCGAATTGTGGCAGGTGCAGTAGTTAAAGAAAA
TCACACGAAATATATTGTATTCTCTCAACATCATGAAGAATATGGTAATAT
45 CGATGATGTACACTTATCTGAACAATATATACTAATGGATCATTTATTATTGA
TACTGATGATTTTGGAGAAGTAGAAAAGTATGATTCTAGAGTAAAAGTATAT
GCAATCACAATAAATAATGATAAAATAAAATTAATAATAGCAACTAAAAT
GAGCACAAGGGATTTATGTCTAATTTGAACTATCAAAGTGATAATATACAT
TTATTGTTAATATGAATGTTAATCTTTGTTAATGAAATTAAGCATAAAT
50 CTCTTGGTGTATTTAAACGATAAATTTAAGTGGATGAGAGGGATTTCGTTT
TGCAAGATTTTGATAACTTAATTCCTGGCTGGTTTAAACATTTGTTCAAG
TCGGGAATGACTTAATTTGGTCTCAATATCTTATTTGGATTATTATTAACAG
CAGGTTTTTTCTTTACAATTAGTTCTAAATTTATTCAACTCAGAATGTTAC
CAGAGATGTTTAGAGCATTAACGAAAAGCCAGAACTTTAAGTAGTGGTG
55 AGAAGGGTATTTACCATTTCAGCTTTTGGCATTAGTGCTGGGTCAAGAG
TAGGAACTGGAAATATTGCCGGTGTGCAACTGCTATTGTTCTTGGTGGCC
CCGGTGCAGTCTTCTGGATGTGGATTATTGCTTTTATTGGTGCAGCTAGTG
CATTTATGGAAGCAACGCTTGCTCAAGTTTATAAGGTACATGACAAAGAAG
GTGGATTCCGTGGCGGACCAGCCTATTACATAACAAAAGGGCTAAACCAA

5 AATGGCTTGAATTGTATTTGCTGTTTTAATTACAGTTACATTTGCTTTTG
TATTTAATACTGTTCAAGCGAATACAATTGCTGAATCATTAAATACACAAT
ACAATATTAGCCCGTAATTACTGGAATAGTACTTGCAGTTATTACAGGTA
TTATCATCTTTGGTGGTGTTCGTAGCATAGCTACACTATCTTCACTTATTG
10 TGCCTATTATGGCTATTGTTTATATAGGTATGGTTTAAATCATTTTATTAC
TCAATATAGATCAAATTGTACCTATGATTGGCACTATTATAAAAGTGCAT
TCGGAGTTCAGCAGGTTACTGGTGGTGTGTAGGAGCTGCTATTCTTCAAG
GTATTAAACGTGGTTTATTCTCAAACGAAGCTGGTATGGGATCTGCACCTA
ATGCTGCTGCTACATCTGCTGTGCCCCATCCCGTTAAACAAGGTTTAAATC
15 AATCATTAGGTGTATTCTTTGACACTATGCTTGTGTGTACAGCTACAGCAA
TTATGATTTTATTATATTCTTGGTGTGCAATTTGGTGATAGCGCGCTCAAG
GTGTAGCAGTTACGCAATCAGCGTTGAACGAACATTAGGTTTACAGCAGGAG
GTATTTTCTTAACTGTAGCAGTTACCTTATTGTGCAATTTTCATCTGTTGTAG
GTAACCTATTACTATGGACAATCCAATATTGAATTTTATCTAACAATAAGA
20 TGATATTATTATTTTATAGATGTTTGTAGTACTTTTAGTATTTGTAGGTG
CTGTTGCTAAACAGAAACAGTTTGGAGTACTGCCGATTTATTTATGGGTC
TTATGGCAATAGTAAATATCATATCAATTATAGGTTTGTGCAATATTGCGT
TTGCAGTGATGAAAGATTATCAAAGACAG

20

Sequence 3393

step.1002e02.cons.ok

25 TGGAGCTGGTAATTTATATATGACGCCACATGATATGGGCAAATTAATTTA
TACGTTACAACAAAAATAAAATCTTTAATGCACGTCAAACCTCGACCTATTTT
ACATGAATTTTGGAACTCAAGAATATCCAGAAGAATATAGATATGGTTTATA
CATAACTCCGTATTTAAATAGAGTCAACGGGGTATTCTTTGGTCAAATTTT
TACTGTTTACTTTAATGATCGGTATATTGTCAATTTTAGGGACGAATGTAAG
TAATACACCTGGATTAGTGAGTAATGAAGACAAAATGAGACACATTTTCTA
TAATATTTCTTGACCAGAAAAGCCTTATAATACAGCAGGTGTTAAAGTTGA
30 GTAATTGTAATTTTGTAAATAATGGATTAGACTCGTTATTCCATATTTACA
AGGTAGTCTTCAATAAATTATATATAACGAATGCCCGGAACCTCAATCTG
AAGTGTAACCGGCATTTGTTATATGATTAATAAAAATAACATTCAAATTGC
AATTTAATTATTTACAATAAGAAAGGGCGTTTAAAGAATAAATCCTCATACT
GTCTTTCAACAGGACTTAAAACAACGTACAATTGGGATGCTCCATATTAA
35 AGTCTACCACGTGGCTCCCCCTTGTAAAGTTGCCATTTAATTTCTATTGAG
TATACGCCTCCTTGACATTAAAGGCTTCTGGATCCAACAAATATAATATAG
TGAATACATCGTATAATTTAAATCCTATTTCAAATCTTCACTCTTATAGT
GCTGAAATAAATTATATAACATGTTTGAAGTTGCATTTGTATCTTTGAAGT
CTTTTACAAAATGGTGAGTAAACAATGCTTCACGAGCCAAATCAAGACCAA
40 TCATAGTCAAAGGTAATCCAGAGTTAAATACAATTTGAGCAGCTTCTGGAT
CACAATATATATTAAATTCAGCTAAAGGCGTTACATTACCTCTACCGGTAC
TACCACCCATTAAAACAATTTCCCTTAATAAATGGTTGAACCTCGGGATAAC
TAGTTAAAGAATAGCGATATTTGTTAGTGGACCTATGGCAATCAAGGTTA
AAGGTTCTTGAGTATTTACTAATAGATTTCTCATTGCTTCAACTGCATGAA
45 TTGATGTTAAATCATCTTGATTTATTTGTGGAACTCGTAACCATCCATAC
CAGACTACCATGAATTGATGTAGCTTCAAAGATGTCATTAATCAATGGTT
GGGATGCCCCCTATGTACAGGAACAGAACTATTAAAAACCTTTTTAGCT
TTAATGCATTTGCTGTGCTTTTCTCAATACCTACATTACCATTACAGTTG
ATATCATTTTAAACGTCAAATTGAGGGTGTGAAAGTGCGATACTAATTGCTG
50 TAGCATCATCTATACCAGGATCAGTATCAATTATAATGGGAATAGACATAT
CATCATCTCCGTATTTTATAATAATTAAAGCATAACATGGTATGGGATTGT
TACGATAACGAAATAAAAAGAGCCTGAGACTTTTAAAGTCTCAGGCTCTTTA
CATATAAGGTGGAATGAATCAATAAAGAACAAATATATTCAAATTCATTC
TTTATTGAAATCAACCCTGGAAAACCTCATCTTTATAAGTGAGAAGAGTGTC
55 CACCTTGCAATTACCCAGTATGTTCCGATAACAGTTACTAAAGTAATGATAA
TTGCAAAGATAACTTTGAACGATTGTAAACGTCCATCTTTACCTTCAGTTA
AATGCATGAACATTAATAATTGAAGGGCAGCTTGAATGAAAGCAAACCAA
AGATGATAGTTACCTTAGCATGGAATGTCATATTAGTGATAGAGTTACAA
AAACTGCTAAAAGCGTTAGTACGATAGAAGCAATAAAACCTACTGTATGTT

TTACGATTGTATTTCATCCGCTATACACCATCCCTATCATATATACGGCAGT
 AAAGATGAAGACCCAAACAACATCTAAGAAGTGCCAGTATAAACTTACTAT
 AAATAATTTAGGAGCATTGTATGAATCCAAACCAGAGTGCCGATTTGAAT
 TAACAAACAAATAACCCAAACAATACCTAATGATACGTGTGCACCGTGCGT
 5 ACCTAGTAGTATAAAGAACTAGACCAGAAGGAGCCAATAGTTGGGTAAAC
 ACCTTCAGAAGCATAGTGTGCGAATTTCGTAAATTTGAAACCTACGAATAC
 AAGACCTAGGATAACTGTGATAATCATCCAAACATCATTAAAGTTTTGTTT
 TTCTTGTCGCATGTAATAAATTGCAATACCACAAGTATAAGAACTAATTAA
 TAATGCAAAAGTCATTATTAAAATCAAATGTAATTCGAATAAGTCGGTAGT
 10 TAATTTACCGCCATATCCGCCACCATGTTGTAACGTTAATAACGTTGCAAA
 TAGGGTACCGAATAACGCAAAATTCAGCTGTAAGGAAAATCCAAAGCCAAG
 TTTATTTAATTCGCCCTTCGTGTGTACGAGAATCAATAGTATTTGCATCATG
 ACTCATGACTTACAGCCTCCCTTTCTTTAATTCGAGCTTCTCTTAAACGAG
 CTTTCAGTTTCTGCAACTTCTGAAGCAGGGATGTGGTAACCATGATCAATTT
 15 GGAACTTCTCCAAATCATAGTAATGAAGATACCTGCTAAACAGATAAGTG
 CTGGAACAATAGATTTCGAAGATTAAGAAGAAACCACCAATAGTCATAAATA
 TACCCATCCAGAATCCTACTGGAGTATTGTTTGGCATATGAATATCTTTGT
 AGTTATGGTTGTCTAAATAATGACGACCATGTTCTTTTCATATCAACGAATG
 TATCGTAGTCAATTCCTCAATCAGGAGTGATAGCAAAGTTGTATTTAGGTGGAA
 20 TAGCTGATGCTGTAGACCATTCTAAAGTACGACCAAGTCCATCCCAGTTAT
 CTCCAGTAGCTTCACGTGGAGCTTTGATATGACTATAAACGATACCTGCAA
 CTAGGAATAAGAATCCAATTGCCATCAATACTGCACCGATAGTTGAGATGA
 AGTTTAGTAACCAACCAACCATCAGAAGGCATGTAAGTGATAGACGACGTG
 GCATACCATCTAAACCTAGAATGAATTTGTGGTAAGAAACAAACGTTAAATC
 25 CGATCATGAAGAACCAGAAGCACCATTGTTTAAATGTTTCATTTAACTTGT
 AGCCCATCAATTTTGGATACCAAGATTAAACCAGCTAAGCAGGCAATA
 CTACACCAGTAACCAATGTATAGTGGAAAGTGAGCTACTAAGAAATAAGTGT
 TGTGATATTGATAGTCAGCTGATGCCATTGCAAGCATTACACCAGTAACCC
 CTCCTAATAAGAAGTTAGGGATGAATGCTAATGAGAATAGCATAGGTGACT
 30 CAAATGTAATTCACCTTTGTATAATGTGAGCAACCAGTTAAATAGTTTAA
 CTCGGTTTGGAACACCGATTAAACATTGTTGAGATAGAGAAGAATGAGTTAA
 TTAACGCACCAATTACCCATAGTGAAGAAATGGTGAACCCAAACTAAGAAAC
 TTAAGAATGCGATACCTGCAGTTGCCCAAATCATACTTTGATGACCGAATA
 AACGTTTACGGGCAAAAGTAGGGATGATTTCTGAGTACATACCGAATGCTG
 35 GCAAAATAACGATATAAAA

Sequence 3394

step.1002e03.cons.ok

40 GCATTAGCACTTAAATGGATAACAAGCCAAATATTGCTACCGCAACAGTT
 GGAGAAGGCAGTTCAAATCAAGGTGACTTTCACGAAGGTATGAACCTTGCT
 GCAGTTCACAAATTACCTTTTCGTCTGTGTAATAATTAACAATAAATATGCG
 ATATCTGTACCAGATTCACTACAATATGCTGCTGAAAAGTTATCAGATCCGT
 GCATTAGGTTACGGTATGCATGGAATACAGGTAGATGGAAATGACCCAAAT
 45 GCAGTATACAAAGCGATGAAAGAAGCAAGAGAACGAGCGCTAGCAGGTGAA
 GGTCCAACATTGATAGAAAGCTGTCACTTCACGTATGACACCACATTCATCT
 GATGATGATGATACATATCGTACAAAAGAAGAAAGAGACCTATTGAAACAA
 GAGGATTGTAATATAAAATTTAAACGGCCTTACTCGATCAAGGCATCATA
 AACGAAAATTGGTTGAGTCAATTGGAAAAAGAGCATAAAGAACTCATTAAT
 50 GAAGCTACTAAATCTGCTGAAGCAGCACCATATCCTTCAGAAGAAGAAGCT
 TTGACATATGTTTATGAAGAGGGAGGTCAACGAAATGACTAAATTATCATA
 CTTAGAAGCTATACAAAATGCACAAGACTTAGCACTAAATCATTTTAGTAA
 TGCATTTATACTTGGCGAAGATGTAGGAAAAAAGGGTGGCGTTTTTCGGCAC
 AACCAAAGGATTACAAAGTAAGTATGGTGATGAACGTGTAATTGATACACC
 55 TCTTGCGGAATCGAATATTATTGGTACTGCTATTGGTGCAGCAATGCTAGG
 TAAGAGACCTATTGCCGAAATACAATTTGCAGATTTCAATTTGCCTGCTAC
 AAATCAAATTTAAGTGAAGCAGCTAAAATGAGATACCGTTCAAATAATGA
 CTGGAATTGCCCCTGACTATCAGAGCACCTTTTGGTGGTGGAGTTCATGG
 TGGATTATATCATTCACAAAGTGTTGAAAGCATTTTTGCTTCAACTCCTGG

ATTAAC TATTGTTATTCC TTCATCACCTTATGATGCTAAAGGTCTTCTATT
 GTCCCTCTATAGAGTCTAACGATCCAGTCTTATACTTTGAACATAAAAAAGC
 ATATCGTTTCTAAAAGAGGAAGTTCCCGAAACATATTATACTGTACCTCT
 AGGTAAAGCAGATGTTAAAAGGCCAGGCGAGGACATCACTGTATTTTGTTA
 5 CGGATTGATGGTGAATTACTGTTTACAAGCTGCAGATATTTTGGCAAATGA
 CGGCATCGATGTTGAAGTAGTCGACTTAAGAACAGTTTATCCACTAGATAA
 AGCAACTATCATTGAACGCTCTCAACGTACTGGTAAAGTTCTTCTTGTAC
 TGAAGATAATCTAGAGGGAAGCATTATGCTCTGAAGTATCTGCAATATATAGC
 TGAAAAC TGTCTGTTTCGATTTAGATGCGCCAATCATGCGATTAGCTGCACC
 10 GGATGTCCTCATCTATGCCATTTTCACCAACATTAGAAAATGAAATTATGAT
 GAACCCAGAAAAGATACAGGACAAAATGCGTGAAC TCGCACAAATTTTAAGG
 AGGGTGTCAATTGGATATAAAAATGCCTAAGCTTGGTGAAAGTGTGCATGA
 AGGTACGATTGAACAATGGTTAGTATCAGTAGGAGATCATGTAGATGAGTA
 TGAACCATTATGTGAAGTTATTACAGATAAAGTAACAGCTGAAGTGCCTTC
 15 AACAA TTTCTGGAACAATAACAGAATTAGTGGTTGAAGAAGGACAACTGT
 CAATATTAACACGGTGAATTGTAAAATCGATTGCGAAAATGGTCAAAATCA
 AACAGAATCGGCAAAATGAGTTTAAAGGAAGAACAAAATCAGCATTCTCAATC
 AAATATAAACGTTGCACAATTGCAAAAATAATCC TAAAACTCATGAAAGTGA
 GGTGCATACAGCCTCTAGTCGCGCAATAACAATGGACGATTTTCACCAGT
 20 TGTCTTTAAATTAGCTTCTGAACATGATATTGATTAAACACAAGTCAAAGG
 AACTGGTTTGAAGGTCGTGTTACTAAGAAAGATATTCAAAATATTATTA
 CAATCCAAACGATCAAGAAAAAGAGAAAGAATTTAAACAAACAGATAAAAA
 AGATCATTCAACGAACCATTGTGACTTTTTACATCAATCCTCAACTAAAAA
 CGAACACTCACCATTATCAAAATGAACGTGTCGTACCAGTTAAAGGTATTAG
 25 AAAAGCTATCGCACAAAATATGGTTACTAGTGTGTCAGCGAAATACCCACGCG
 TTGGATGATGGTTGAAGCTGATGCAACGAATTTGGTTCAGACTAGAACTA
 TCATAAAGCTCAATTTAAACAGAATGAGGGTTACAATTTAACTTTCTTTGC
 GTTTTTTGTAAAAGCTGTTGTCAGAGGCTTTAAAGTAAATCCATTACTCAA
 TAGTACATGGCAAGGAGATGAAATTGTTATCCACAAAGATATTAATATCTC
 30 TATTGCTGTTGCAGACGATGATAAGTTGTATGTGCCAGTCATTAAAAATGC
 AGATGAAAAATCAATTAAAGGTATCGCGCGTGAATCAATGATTTAGCTAC
 TAAAGCAAGATTAGGAAAATTAGCACAAAGTGATATGCAAAACGGTACATT
 TACGGTTAATAATACTGGTTCTTTTGGTTCTGTTTCTTCAATGGGAATCAT
 TAATCATCCACAAGCTGCCATTTTACAAGTAGAATCAGTCGTTAAGAAACC
 35 TGTAAGTTATAGATGATATGATTGCAATTAGAAATATGGTTAATTTGTGTAT
 TTCAATCGATCATCGTATTTCTCGATGGTGTTCAAACGGGAAAAATTTATGAA
 TCTTGTTAAGAAAAAATAGAACAATATTCTATTGAAAACACTTCTATTTA
 TTAATTACAAACAATGATTACCAATTAAACTTCAATTAAAGTTGAACGCT
 TTAAGGAGGATAATTTCTTCTTAAAGCGTTCTTCAGTATTCGTTTAAATT
 40 TCAACTAGAATTTTTCAATCATCATAAATTAATCCTGTTGAACAATGTGTT
 GTTAATTAGTAGAATAAAATAAGTAAATTAACCTATTATGGAAGGTGACA
 AATCATGGATTTAACTTTGATTTATATATGAATGATGTTGTTGAACAAGC
 ACGAAACGAAATTGAGCATGCTGGTTATCATCAATTAACCTCTGCTGAGGA
 TGTTGATCAAGTTTTACAACAAAAGGGGACATCTTTAGTCATGGTAAATTC
 45 CGTA

Sequence 3395

step.1002e04.cons.ok

50 AGACCGTCTACAATATACAGTTTTTACTGTTGAAACACACACTACTTTTTT
 ACATGAATTAACACTTGGTCAAGAATTCAATATAGAGCTATACCTTTATAA
 TTACGATGACAAACGAAC TCAATTTCTTTCTGAGAATGTTAATTGATAATCA
 AGAAGTTGTTGCAACCAATGAGGTCATGATGTTAGGAATAGATAGGACACA
 ACGTCGTGCTGCACCATTTC CAAAACATTATCTGAATGCTATACAAGATTA
 55 TGCTCATAAGCAAGAAAAAATGAATGGCCACCACAATTAGGCCACTCAAT
 TGGAAATACCATATAAAGGAGAATAATAATGTC TATAAAAGAAAAATTAGTA
 ATAGAGCTCATGAATTTAAAGACACTG TAAAATACGAATTGTCAGAATTA
 AATGAACAACAATATTTGTTTATGAACGGCC CAGCACCTCAGTTGATGAAG
 CGTGCCATGAAAATGAGCTATATTTCTAGGCC AAAAGAAGCGATTGATCAC

TTTCTATTACTCATCGACACTTGCCATGAGAATGTCCTTACTAGATAACTGC
 CAAGACTATCAAAAACAAATTC AACATAAACAAATTCCTTTACAAGAAGAT
 ATTATTACCAAATAAATCAGTCAAAGGAATTCGAATCATTTTTATCAACT
 TATTATGTAAATAAAGGTAAATATGATATCACCAGCAAAATAAAAACATTA
 5 CTTGAATAAAATACTATTGTTTAAATAAAATGATTATCTCATACATCACTAG
 CTGATAAGCATAGCATTGGTAATTGACTTCATCCTTCCTTTTAGAACAAACA
 TTCACATCACTTTTAAATACAGTTAAAAATCTATTCAATTCTTACTCTAA
 GATTGATTTAGTGATGTGTTTAAACACCTTGTGAGATTGGCGTATGGGATA
 GATAGGAAAAATCATGCACCAATTTTAGGATACTCTTTAAACTCAACATCGTT
 10 CCCACATTCCTCTAGAGCATGAGTAAATGCATGCATGTCTGGGCTTAATAT
 TTCTCTGCCACCTCCATACATATATATAGGTGGCAACCTTTTAGAGTACC
 GTATAATGGTGATATACGCGCTCTGATAAAGGTAAATCATTTCGTCCAAGA
 TTTTCATGAGTTGATGAACACCAATCGACTAACTAAGATATCATTTCCTTC
 CAACGTTTTTCGTAATATTAGGATTAGTCAAAGTGGCATCTAAAAGCGGTGA
 15 AAGTAAAAATAACTTTCTAGGCACTTCTTGATTATCATTGATAAGTGATTG
 AACAAAATTAATGCTAGTCCCTCCACCAGAACCATCCCCATCATAACAAT
 GTTGTTCGCACCAACTTCTTCAACTAATTGATTATAAACGTCACGAATCGC
 TTTGAACGTTTTCAAGTAGTGATAATCTGGTGACTTTGGATAGATAGGTAA
 AACGACCTCATGCAATGTATTTAAAGTTAGTTTATCTAAAAGTCTCCAATG
 20 AAATGGAGAAGGTTGCAAAGTATTATATCCACCATGCAGATAGAGTATCTT
 TTGATTTTTCTCATGTCTAAAAATAAATCTAAACACTTGCATATCATTAAG
 AGTAAGTTTTATCTAAATTTGACTTTACATTAAAGTGTGCGAGGTTGCTGATG
 TTTTTTACTATTTTCGATACTTCTCTTTTCTAAAAAATGTTTAACTTCATC
 ATCATTACTAAAAAATATCGATCTATTGTGAAGTATATATTTATTAACCAC
 25 CGGATTCATCATTTCTATTTTTTCATTAATCAATCAACCTACCTCTATTTAAA
 ATTTTAAATTTTAAAGAAATAATTTGTTTAAATATCTATATACCTTTGCTAAA
 AAAAGATATAATGATAAGTATGTCATTTTAAAGAAATATTTTTTAATCAAT
 ACTTAGGATACTAGATTAAACCAATTATAACAGACATAAAATTA AAAAGA
 TAAACTTAGTAAATTTGCATCTAAAGGAGGCATTTAAATGTCTGAAACATT
 30 ACAAAACCAAGACAATTCAACTCTCAATTTGAACATCAAGAGATCCATCG
 TGGAAAAAGATATGGTAAAAAGAAACGCTCATGGGTAAGTCTCATTTATCA
 AGTTTTTCGTTTTAGTATTAACTGCTATTACTGGCTATAGTATGCTTAAACA
 ACCTATATTTAAAAATTTCAATTTGTAAATGAACTATAAAATTTCCATCAATT
 AAGAAATTTTCAAGATACAGTTACGCAAATCGGTAATTTGAATTTAGGTAA
 35 TATTGATCAATTACAACAGTCTGTCGATAATCTCATCATTATATTTAATAT
 TTTCTTTGTCTATGTCTGATTAGTTTGTTTATAACCGTAATTACAATTAT
 TTTTAATCGCACAGCACTCAAAGTAGTCAATATATTACTTTTTAGCTATTA
 TGCTAGTCAATTACATTGTATTTTCAAGTTATATCATACATACAATTGCTCAGA
 AAATTTCCGAACTCTCTAAACAATATTATTTAACAGTATCACCAGAACAAAG
 40 TATTAACGAAGCAGATGCGATTCACAATGCACTGATACTCATTTGGATGTA
 GTATTGCGTTATTGATTGTAAGTTTGTTTTTCCGTAATCGTTTACCACGTA
 TTAAATAAAGATGATTTTAATGTTCTTTTAA

45 Sequence 3396

step.1002e05.cons.ok

CCATAGCAACATTTTCAGTTATGAATTGATTACTTTGTATGGGATGTGTAA
 CTTCTATCACTTTCATAATATTATTTCTCCTTTAAATAAAAACTTTTTAG
 GTTTTATTTTCATAAGGTTTATCTGGATGAACATATATATTGCTAAAACTT
 50 TTTGAGTGCTACTATCGACAAATATAAAAGGTTCTTTAACATTTTGACTTA
 ACACTTTTTTATGAAATTTTTGACCGTTACAGATTTCTTTTTAAAAATTTG
 AATCTTTCACTTGGAAATGATTTCAGACCTTTTAAAGCCATATTCTATAGGAA
 ACAATTCATTATGTAATGAATCATGCTCATGTAATCTTTAATTTGATCAA
 TCGTTAAACTACTTTCTAATTTGAAAACCGCCAGAAGCAATTCTAGTTAGAC
 55 GTGACATATGAGCTGGAAAACCAAGTTTAAAGTCCAATATCTGTAGCTAAAG
 TTCTAATATAAGTTCTTTTACCACATGTTACTTCAACTTCAAAATGACATG
 TCTGCTCCTGAAAAGTAACCTCAGATATTCTATGTATATCTTTAATAAAAA
 CTTGTCTGTTAGGGCGTTTCGACAGTTTCATTATTTCTCGCATATTCTATA
 ATTTTCTTCCATTTACTTTAACAGAAGAATACATAGGCGGAATTTGTTGAA

TATGCCCCCTCAAATTGCTGCAACACTTGGTCAATCGTATCTTCATTAATAT
CATTCTTATCAACAGCCCTAGTTTCTAAAAATATCTCCAGTTTGGTCTTCAG
TAGTCGTACTCTTTCTAGCGTTATCATAGCGTGATAAGTTTTCCCATTT
CCATGATATAATCACTGACTTTTGTGCGATCGCCTAAACAAATTGGTAACA
5 CACCATTAACTTCAGGATCTAATGTTCCCTGTATGACCAATTTTTTCATTT
TTAAAAATTTTACGTAATTTAAAAACGACGTCGTGACTTGTTAAACCTCGTT
TCTTAAATACCGGTAGTATGCCATTATACATGTTGTCACACCTTTAACTTT
TAATAACGTGTATTATTATAACACGATATTTTAATTCTTTAATTAGCGTTT
TACATAAAAAATAGTAGGCAATATCAAAAAATATGCTTTATGAATGATTGTGC
10 CTACTATAAAATATATTATTTATCATTTTTGTGTAACCTCTGAATCATGCG
TTCTATCTTATTACCGTATTCGATAGATTTCATCATATTCGAATGTTAACTC
AGGTATAAATCTTAGGCGCATTTCGAGAACCAAGTTCAGATTTTATAAACCC
AGTTGCTTTATGCAAGCTTTAAACGTATTATCAACTTCTTTATCATTCCC
TAACACTGTTAAATATACCTTTGCTTGTGAAAGGTCATTGGTTAGTTCAAC
15 ATCAGTAATTGTTAAAAAACCAACTCTAGGGTCTTTAACTTTATTATTAAC
AATGTCCATGATTTCCTGTTTCATTTGTTCTCTACACGTTCTGCTCTTAT
ATTATTCATCATCTCACCTCTCTTTATTTTACTTTACTCACAAATAAATG
TGAAATATCAGATTAAATGATTAACTATTGATGTTTCATTTTATTAATATAT
CAACAACGAATCAATAAAATAATACTAAATGAGAGACCATAACTAAAGTAT
20 TGATAGCATAACATTAAGCTTTTCCAAATCAAAAAATTTATATCTTCAATTT
GATTGTTGGTTAAGCAAAGTATGTCATACTCTCCTTTAAGATATTAGGTTT
ATGAAATATCTTAAAGTACTTAACTTAGCATATAGCCTCTCAATAAGTTTC
AAAAATTATCTTTGAATTTCTACCATTTCAAACGCTTCAATAATGTCCTCT
TCTTTGAGATCATTATATTTTCAATTGTAATACCACATTCATAGCCTTGA
25 GCTACTTCTTTAGCATCATCTTTGAAACGTTTTAATGTGTCAAGTTCACCT
TCAAATAACACGATACCATCTCTAATTACGCGTACACCAGCGTTACGAGTG
ATTTTACCTTCAGTCACATAACTACCAGCAATTGTACCAACTTTAGAACT
TTAAATGTTTGAACGACTTCAGCTTGTCCAATGACTTGCTCTTCAAATCT
GGGTCAAGTAAACCTTTTCATAGCTGATTCTATCTCTTCAATAACATTATAG
30 ATAACCTCTGTGTAATCGCATATCTACATTTTCAGCTTCAGCCGCACGTTTC
GCACCTGCATCTGGGCGTACATTAAACCAATAATAATACCATTTGATGCA
TTTGCTAATGTAACATCTGATTTCATTGATAGCACCAACAGCTGTATGAATA
ATTCGTACATTACACCTTCAACATCTATTTTCATTAGAGATGCGGCCAAT
GCTTCAACTGAACCTTGTACATCACCTTTAATGATGACATTTAAATCTTTC
35 ATTTACCTTGTGTTTCATTTGCTCAAATAAATTGCTAATGAAACATTTTTC
CTTTCTTGACGTTGCTGTATGACACTTGCCCTCATGACGTGCTTCACCAATT
CGACGTGCTTGTTTTCATCACCAAAATACAACAAAACGATCACCTGCAAGT
GGAACATCGTTAATACAGTAATTTCTACAGGTGTTGAAGGACCGGCAGAT
TTAATTCTTTTTCTAAATCATTTAACCATTCACGATACGTCCTAAGTA
40 TTACCTACAACAATTGCATCTCCAACGTTTAAAGTACCATTTTGAACAAGT
AAAGATGCAGCTGGACCTCGTGATTATCTAATTCAGCCTCAATCACAGTA
CCTACAGCTTGTATTAGGATTAGCTTTAAGTTCTTGTACCTCCGCTACT
AAACCGATCATTTCTAATAAATCATCAATACCGTCTCCACTCAATGCAGAT
AGTGGTACAAAGATTGTGTCACCGCCCCAGTCTTCTGGAATTAATCCATAC
45 TCAGTGAGTTCTTGCATAACACGATCAGGGTTAGCAGTTGGTTTTATCAATT
TTGTTTACTGCAACAATCGTAGGTACTTCTGCTTCTTTAGCGTGATTATA
GCTTCAATTGTTTGAAGCATCACACCATCATCAGCGGCCACGACTAAAAAT
GTAATATCAGTAACTTGAGCACCACGTGCACGCATAGTCGTAAATGCAGCA
TGTCCAGGAGTATCTAAGAACGTAATTTTTTTACCTGAATTTTCAATTTGA
50 TAAGCACCAATATGTTGAGTGATTCCGCCAGCTTCTCCTTCTGTAACCTTTA
GTGTTACGAATAGAATCTAATAAAGTCGTTTTACCATGGTCTACGTGGCCC
ATGATTGTAACAACTGCTGGACGTTCAATTGCATCAGAATCATCAGTCTCA
TCATCAAAATAAATTGATAAATCTTCTTCATCAACGACTACTTCTTTCTCT
ATTTCAACGCCATAGTCATCTGCAATTAATTCTAATGTTTCTTCATCCAAT
55 GATTGATTGATATTAGCCATAATACCTAGTAAGAACAATTTTTTAATAATA
CCAGCTGATTCTACATTTAGCTTTTCAGCTAACTCACCGACAGTTATGCCT
TCTTGATAAGTGATTTTAGAGGGCATTTCTTTTGTTCGTACCTCATTTT
TGAGGTTTATTATTTTATTATCTTTTATTTTATGTTTTTTTGATTCT
TTATTAGTTTTATTGTTTTATTGTTGCTGTTTCCTTGTTGTTTTGTTCT

TTTTTCCTCGTTGGTTTACTATTATTCTTACTTTGTGTTTTCTTTATCG
TTAGAATTTGTTTTATTATTAGATTTTTGGTGATTATTTGAGTATTTGT
TTATTAGTGCTTTTCGCT

5

Sequence 3397

step.1002e06.cons.ok

AAAATTCAATATACGTTTCTTTTATAATTTTAGTAATTTCTCTTTAAAAAT
ATCACGTAAAAATTCAGGCATAATATAATCTGTAGATAACCCCTTTAGCAA
10 TGGGAAATATCTTCCAAATGTATACATATCAACAGTACCTAACGTATATTG
TTTATCATTTTGAATCTCTTCATCGCCTATAAAATAACGGTCTTCTGACTC
AATAAAGCCTAGATTATATAAAATATATCCTCCAAATATATCCCCCTTAAA
ACCAAGACCTTGTGCATGTTCTGTCATATATTCTTGCTTTTGGCTTTTGTG
AATCACTTGCTTTAATTGTTTACCGGTTAATCGAACTCTTACAATATTGAT
15 TGGATGGGGTAAACATGCGATGTATATCATATTCCGTCACCTTTATCAGCTTC
AATGCCATTAACTATAAGTCCAGCATTTACGATTGCACAATCAGCCCTTGA
AAACTCATATACACTTTCAGCCAGTAAATACGATGTTCTTGTAACAACATC
TGTTCTTTTGAATCTTCAAGTTCACATGATGATTAAGTACTGGTTTACTTAGAAG
TGCTCTTCTTCTTCTTCAAAATGTGTCTCGACTAAGGGAAGTGTTCAT
20 AGGATGAATTTTGGCGATTTTATCAACGATTTTTCATTTTCAATCGTAAT
ATTAACCTTACCTAAATAATAGCCATATTTTCCGGCAGCTGCCATCAAAAC
ACCATTGTTTATTTCTCCATGTTCAAAATGATGATGCGTATGACTACCAAA
GATAACATCTATTTCCGGAATCTCTTGGCATAAATTTTCATCAAAAAAGAT
ACCGACATGGCTCATAACCATTAAGATCATATTACCTTGATGTGCATT
25 GATTTTCATCTTTGATTGCGCTAATGGGTCAAGTAAATCCAATCCAGTGC
TCGATAAAAGGTGTGAACGGTGCCGTTGCTGCAACAAATAAAATACGTGT
TCCTTTTATTTCTTTGATATACGAAGAGGTAATATGATGTGGAAGATGTCC
CTCTTCATCTATGACATTCTGTGCAATCACTTTAAATCCGCGTCGTTATA
TAGATTTTGTAAAGCATCATGAGAAATGTCAATTCCTTCATTATTTCCAAT
30 GGTTGCAATATCACAATGTGCTTCATTTAAAAGTCTATATTTTATGTCC
TACCGTAGCTTCTGTACAGGTGCTGATAAATCAACATGGTCACCTATATC
TATATAGAGTGAGGGATGTTCAAGTTGCGGTCTATGTTTTGCCATATAAGC
TTGAATACGAGCATATTCAATTTAAATGACTATGAATATCATTCTGTGATA
TATTGTTAATCCATAACATGATCCCCCTCTTTAAATTTATATACCTAAT
35 TGTACTTTAGCATTAAGAACAAGTCTCAAAGCATAACTACTTCGCATTT
AAATAAAAGATTTAATGATTAAATATACACCCATGATAAGCATTACTGTTC
TCAATAACATTACAACGTGTGTCGGATTTTCATAGATCGATTGACCCTCACAC
CTATTTGTGCACCTATAACACTTGAAATAATGAGAATGATAGAATAGCCCC
AAGCCACATGTCTTGAAGATGTGCCCTATTGAACTCATCACACTTGAAA
40 AGAAAATCATCATCACTTGTGCTTACTGCAACATGTGGTGGAATCTAA
AGACGATGAGCATAAGAGGGTCAATCAATGCACCTCCACCTATGCCAAACA
GTCTGTTTAATAGCCCAATAAATAACGTTGTTATAAAAGCAAACAGGGGTG
GTACACTATAACGGTATGTTTTCCCTTCAGCATCAACATATGTACGCGCAT
ATTTAGGTTTGTCAAAAATTTTAAAAGGTTTGATTTGTGACGAATCATCA
45 GTAATATCGACACAAATATCATAAACACGCCAAAGTATAAGTTAAACGAAT
CAAGTGTAAGGTATCTACTTAAAAAAGAACCAACCAAGATCCAGGTAGTA
GTCCAAATAGAAATATCGAACCATTTTTATATCTACTTGTGTTGGTTTATA
AATATCCTAATGAGGAAGATAATCCAGTAACAATCAGAATAACTGAAGAAG
TACCTATTGCAATTTGCGTCGTAATACCATGTAGAAGATGATGGTTAACGC
50 CTAAATAAACAAGTGTGGGACGATGATGATTCCACCACCAATACCAACGA
TAGATCCTAAAAATCGCTGATAATCCCCCTATTAAATTAATAGAATAATTG
TTAATAACATGACATTGCGCTCTTAAAAATAGTTTAAATTGTTGCGGTGCTA
AACCTTCATAATCAATGTCTAAAAATATTTTGATACTTTTTAGCATTTATTAG
CTGCATGACCGCCAGAGTTATTATTAATAAATTACATATACTTTCTTAGCCT
55 TTTGATTAAAGTATTTTCGACTTTACGAGCCAAGTCAGCTAACTCATCATCGC
TATAATCATATAAATATCTTACATCTCGCCATTCTTGATCAGTCATATCTT
TTTTAGTCCAACCATAATGATTACGTCCATGATAACGTACAAAAGCAATTT
CACTAGTAATCCTATTTACTAAAGGAACGCTCCCCCTTTAACTTGAGGTT
CATCTACCACTGCATGAATGATTGATGTTGTGTTAAGAAGGATAAAGTTT

GTTCTTTTATACTGATTGTCAAACCATGATTGATGTCTAAATTCATGCTCA
 TCGGAATATCAGTTAAATTGTTTCTCACATATAAGATGTAATTTATATTTT
 GGACGTTACAATCAAACCAAGGTGGAAATTGTACGAGCACCATCGCCAATT
 TATTATGCTCTTTTAAACGGATTGATCATTCTTTAAACTGTTTAAATAATA
 5 TATCTCTCGACTCTGCATAATCTCTATAGTCAGCATGCAATGTTAATGCTT
 GATGAATTTTAAACAACGAATTGAAATCTATTCGGCGTTTCATTTATCCATT
 TAATTATATTTTCGCTCTGGCTGTATCGCATAATACGAAGCATCCAACCTCTA
 CTATTGGGAAGTGACTAGCATACGTTTAACTTATCTGTTTTACGCTCTA
 AATCCTCATATAATGTATCATGGTCACCCAGCCTTCTAAAACTCTGGAT
 10 AAAAAATCACAATTACCTTTCACACGCTGATACAACTAAGTTCGTGTCAT
 TATAACTAAATTGAGAATCCCTTAAATGTAAAAATTTATCTGCATCTTTAT
 GTAGTTCATAAATATGATATAAGATACCTGTTTCTTTGTGATTATCACTAA
 GAATAATCCACTTAGACATGTTGTTTACCCTCTAAATATTCTTTTAAACAGT
 CGAATAGCATTTCCACGATGACTGATTTTACCCTTTTCATCATTTGTAATT
 15 TCCGCCATCGTTTATTTAATTCTGGGACAAAAAATATTGGATCATAACCA
 AAACCATCTTTCCATGTCGTTTCAGTTGTTATCACTCCAGAAACTGTTT

Sequence 3398

20 step.1002e07.cons.ok
 TTTGGAAGTCTGTGTTGATAAAATCAAATTATTCGCTGAATTGACTTCTAAA
 TATAAAGATTACCATCCTGAGTATCTGGGTAAGTTTAAACAGATATGAAATA
 GATAAATATATTCAAGAAATTAGCAAAAGTTATAAAAAAGAGGTAATGAA
 AATGGCTGAAAAATCAAGCTAAAAAAGAACAAGTTCATACAGTATTCCAAAA
 25 TATTTCTCAAAAAATACGATCGACTCAACAACATCATAAGTTTCGAACAGCA
 TAAAGTATGGAGAAAACATGTTATGAGTACTATGAATGTACAAAAAGGCAG
 TAAAGCTTTAGATGTGTGTTGTTGGTACTGCAGATTGGACTATTGCATTAAG
 CGAAGCAGTAGGCTCAAAAGGCCAAGTAACGGTCTTGATTTTAGTGAAAA
 TATGCTTGAGGTAGGTAACAGAAAACTGCTTCATTGGAAAAATATCCAGCT
 30 TGTTTCATGGTGATGCTATGAATTTACCTTTTGATGATAAATCATTGATTA
 TGTTACTATTGGTTTTGGCTTACGCAATGTTCCGGATTATTTGTCAGCACT
 CAAAGAAATGCATCGAGTTTAAAAACCAGGAGGCATGGTGGTATGCTTGA
 AACAAGCCAAACCGACATTGCCATTATTCAAACAAATCTATAGTTTATACTT
 TAAATTTTGATGCCAATTTTGGTAAATGTTTGCTAAGTCCAAAGAAGA
 35 ATATGAATGGTTGCAACAATCAACATTTAATTTTCTGACAAACAAACATT
 GAAAGGCTTATTCTTCAAGCAGGATTTAACGATATCATTGTACGTAGCTT
 TACAGGTGGCGTTGCTGCAATGCACCTTGGCTATAAAGAAAACAGTTCTAG
 CAAAGGTGATTAACGTGGCAAAGTTAAACATTAAACAACGAAATAAAGAAAG
 TAGAAAAGCGACTTGAAGAAGCAATTATAAGTTCTGATCAAACATTACAAG
 40 AAGCCTCATTCATTTTACTATCTTCAGGGGGAAAAAGAGTTAGACCCGCTT
 TTGTTATTTTAAAGTGGTCAATTTGGCTCTAACAACAAACCTTCAGAAGACA
 CGTATCGTGTAGCAGTAGCTTTAGAACTAATTCACATGGCTACCTTAGTCC
 ACGATGATGTGATAGATAAAAGTGATAAACGTAGAGCCGACTCACTATTT
 CAAAAAATGGGACCAAAAGTACAGCTATTTTAAACAGGAAATTTCTTACTTG
 45 CTATGGGGCTCAAGCATTTATCTGAAATCAGTGATACTCGTGTCCATTTCGA
 CCATTTCTAAATCAATTGTTGATGTGTGTAGAGGAGAATATTCCAATTTT
 AAGATCAATTTAATAGCAATCAAACAATTACTAATTACTTACGTGCTATCA
 ACCGTAAACAGCACTTCTTATTCAACTGTCTACACAAGTTGGTGCGATTA
 CTTCCAATGCGTCAAATGACGTTATTTCGTAAATTTAAAAATGATCGGACATT
 50 ATATAGGTATGAGTTTCCAAATAATAGATGATGTGCTAGATTTTACTAGTT
 CTGAAAGAAACTTGGTAAGCCGGTTGGTAGTGACCTTATGAATGGTCATA
 TTACATTTACCTGTACTATTAGAAATGCGAAAAAATAAGACTTTTAAAGATA
 AAATTTCACAACTTAATCCTGACAGTCCTCAACATGCCTTTGAAACTTGTA
 TAACAATAATTAGACAGTCCGAAAGCATAGAACAATCAAAACAAATAAGTG
 55 AAAAGTATTTAAATAAAGCAATCAATTTAATCGATGAATTAGAGGATGGTC
 CTAATAAAGAACTATTTAGAAAGCTTATTAAAAAAATGGGAAGTCGAAATA
 AGTAAGTTTGTGACAAAAGTTGAAAGCGCTTCATTATCCTGTTACCATAC
 TAAGTAGCAGGATATTTTTATAGCTTGCAATGTTTAAAGTTTTTTACATAC
 TTATAAAAAATGAATGAAACACATGCACAGGGGACACACGAAGTGGAACGC

ACATTTTTGATGATTAAACCGGACGCTGTCCAAAGAAATTTAATTGGTGAG
 ATTATATCTCGCATTGAAAAGAAAGGTCTTAAGCTTGTGGCGGTAAAGTTT
 ATGCAAGTGCCCTATGGAACCTAGCTGAAAAACATTATAGTGAACATGAAGGA
 AAACCGTTCTACGACAAGTTAATTTTCATTTCATTACTTCTGCACCCGTTTTT
 5 GCTATGGTAGTAGAAGGTGAAAAATGCTGTAGCTGTATCTCGTAAAAATTATT
 GGAAGCACAAATCCAAGCGAAGCAGCTCCTGGTACAATCCGAGGCGACTAT
 GGTTTAAATCTAGGTTCGTAATATATACATGGTTCAGATTCAACAGAAATCA
 GCACAACGTGAAGTTAAATTTATGGTTTACAAGTAGCGAAATTTGCTGATTAT
 AAAGAACCAGAGAAGATTGGCTATACGAATAATAAAAAATATCAAATAATG
 10 AAATTCATTTTATCCATTAAATGTCTGTGACGCATATTTGTGTCTCAGGCAT
 TTTTATGATTTAACCAATTTTATTTGCTTAAAAATAACTTATTGAATTA
 GTCTATTTATTTTTTTATCTAGGGATTTAACGATGCAATACCGTATATTGT
 TCGATTATATAGCAATTTTATCTTTTGACAATTGAGCATTCATTATTAAGT
 AATGACCTACTTTTAATATATCAATCATGTATCTAATGATTTTCGGTCTTTT
 15 GAGTTTTTTAATAGACTGATCTTTAATTGGTATAATGTCACATGACATAA
 CAATCATCTTGGTGCTATTCTTTTAAATATTTATAACTGCTAAAAATAA
 AATCTCCTTGTGTGCTAACTGACTTATGTTGTGTGATTGAATATGATAAT
 ATATTAAACATATTACAATTTAATGTATTGAAGAATAGAAAGGGCGTGCG
 ATCATGAGATACCTTACATCCGGTGAATCACATGGACCACAACCTACAGTC
 20 ATCATAGAAGCTTTAGAGCACAGTGGCGA

Sequence 3399

step.1002e08.cons.ok

25 TTTAAATTGTCAGATGATCAAGAAAATGAAATTGAGCAATTATTAGATCAA
 ACCAATCCTGATTTACCACGACCAGTAGGAGAGGATATTGTACATTATTCA
 GATTATTTTGAAGGTGCACAAAAGTATCTAAGTTATCTTAAATCAACTGTT
 GATGTTAAATTTTGAGGGTCTTAAAAATTGTATTAGATGGTGCAAACGGGTCA
 ACTTCTTCTTTAGCCCCATTCCTGTTTGGCGATTTAGAAGCGGATACTGAG
 30 ACAATTGGATGTAATCCAGATGGTTATAACATTAAATGAACAATGTGGCTCT
 ACTCATCCAGAAAAATTAGCTGAAGCTGTGTTAGAAACTGAAAGTGACTTT
 GGTTTAGCTTTTGATGGAGATGGCGATCGAATTATTGCGGTAGATGAAAAT
 GGACAAATGTAGATGGAGATCAAATTATGTTTCATTATTGGTCAAGAGATG
 TATAAAAAACCAAGAACTCAATGGAAATATGATAGTTTCGACAGTAATGAGT
 35 AACCTTGGTTTCTACAAAGCTCTAGAAAAAGAAGGTATTTCAGTCAAACAAA
 ACTAAAGTTGGAGATCGCTATGTTGTGCGAGGAAATGAGAAGAGGAAATAT
 AATCTTGGTGGTGAACAATCCGGTCATATCGTATTAATGGATTACAATACT
 ACTGGTGATGGATTATTAACGGGTGTTTCAGTTGGCTTCCGTTATTAAAAATG
 AGTGGTAAAACTCTAAGCGAGTTAGCTTCTCAAATGAAAAAGTACCCACAA
 40 TCTTTAATTAATGTGAGAGTGACTGACAAATATCGTGTTGAAGAGAATATT
 CATGTTCAAGAGATAATGACGAAAGTTGAAACAGAGATGAATGGTGAAGGA
 AGAATCTTGTTCGTCCTTCTGGAACCTGAACCTTTAGTACGTGTAATGGTT
 GAGGCTGCAACTGACGCGGATGCTGAAAGATATGCTCAAAGTATCGCTGAC
 GTTGTGTAAGACAAAAATGGGCTTAGATAAAATAACTTTTCATGTTTAAATT
 45 CATATATAAAACAGATTTATGATGTTACAGACATAAATCTGTTTPTTAA
 CAGAATTTTATCTTTTAAAAAGATTTGAAAAATATATTTTTTTATTTTGAG
 AATAATTCATCGATGTATCTATTCAGAACTATTACTATTTGTGAAAAATAC
 TGTTAAATCGTTTGCAAATTAATTCGTTGACAGTGATACAATGAATATTAAT
 GGAAAGGAGGAGAGTCTAGGTTATTAATAATTAAAGCGCCTGTGCAAAATA
 50 CCTAAATAGTTATCAAGTATTAAGTTAAATAGTAGAACTTTTATAGAT
 AACGATTTTGGATTTGTAGACGAGGAGGATAGTGATCGAATCAGATCGGCG
 GATGCTATCCCGGATGTGGCACATTTCGTTAGCTTATTAAGTAAATCATTA
 GGTGACTTAGTGACAAAGTTAATAAGATCGCCAACTAAATAAATTTATTT
 TGGGAACTACGGATAGCGTGATAAAATTCATATAATAGCACGCTTATTT
 55 CATGATGTATTTCGTTGATAATGATTATGATAGGGGATTATCAAATTTAAGC
 AGTGATAATTAATCATAATAGCCGTGTTATCCCATTTGTTGGAGGATATTA
 TATGTGTGGAATTGTTGGTTATATTGGCTACGATAATGCCAAAGAATTACT
 ATTAAAGGGTTAGAAAAATTAGAATATCGTGGTTATGACTCAGCAGGTAT
 TGCTGTAGTTAATGATGATGGTACAAAACCTTTTAAAGAAAAGGAAGAAT

5 TGCTGAATTACGTAAAGTTGCAGATAATAGTGATGAGGATGGTACGTTAGG
AATTGGTCATACACGTTGGGCGACACATGGTGTTCCAAATTATGAAAATTC
ACATCCACACCAGTCAACATCTGGACGTTTTACATTAGTTCATAATGGTGT
AATTGAAAATTATGAAGAATTAAAAAGCTGAATATTTATCTGATGTCACTTT
10 TTCATCAGAACTGATACGGAAGTTATTGTACAATTAGTAGATTATTTTTC
TAGACAAGGATTAGCTACAGAAGATGCATTTACAAAAGTAGTTAAATTATT
ACATGGTTCATATGCTTTAGGATTATTAGATGATAATGATAAAGATACTAT
TTATGTGGCTAAAAACAAGTCTCCGCTTTTAGTAGGTGTAGGTGAAGGTTT
CAATGTTATTGCTTCTGATGCTCTAGCAATGTTACAAACTACAAACCAATA
15 CAAAGAGATACATGACCATGAAATAGTTATTGTTAAGCGAGACACAGTAGA
AATTAAAGATCTTGAGGGGCACATTCAACAACGTGATACGTATACGGCAGA
AATAGATGCTGCTGATGCAGAAAAAGGCGTATATGATCATTACATGTTAAA
AGAAATTCATGAACAGCCTGCAGTGATGCGTCGCATTATTCAAGAATATCA
AGATGAAAAGGTAATTTAAAAATCGATTTCAGAGATTATTAATGATGTAGC
20 AGATGCTGATCGTATTTACATCGTTGCAGCTGGTACTAGTTATCATGCTGG
ATTGGTTGGTAAAGAATTTATTGAAAAATGGGCAGGTGTACCTACTGAGGT
TCATGTAGCTTCTGAATTTGTATATAATATGCCACTTCTTTCTGAAAAACC
ACTATTTATTATTATTTTACAAATCTGGTGAAACAGCTGATAGTCGTGCTGT
ATTAGTTGAAACAAATAAGTTAGGTCACAAATCATTAACAATTACTAATGT
25 TGCTGGTTCAACATTATCACGTGAAGCGGATCATACTACTTTTACATGC
TGGACCTGAGATTGCAGTCGCATCTACAAAAGCATATACAGCGCAAAATTGC
TGTTTTATCTATCTTATCTCAAATTTGTTGCTAAAAATCATGGTCGTGAAAC
CGATGTTGATTTATTAAGAGAACTAGCTAAGGTTACTACAGCTATTGAAAC
AATTGTTGACGATGCACCTAAGATGGAGCAAATTGCAACGGATTTCTTAAA
30 AACTACTCGTAAATGCATTCTTCATTGGACGAACAATTGATTATAATGTTAG
TTTAGAAGGTGCATTAAAAATAAAAGAAATTTCTTATATTCAAGCTGAAGG
ATTTGCAGGTGGGGAATTAAAGCACGGAACAATCGCTTTGATTGAAGATGG
CACACCTGTTATAGGTTTAGCTACACAAGAAAACGTTAATCTATCAATTCG
TGGAATATGAAAGAAGTACTTTAGAGCACAGTGGCGATGATATCA

Sequence 3400
step.1002e10.cons.ok
35 TTCCCTCATAACGACCCTTTTCAACTTCTCTTGGATCATATTTTGGTTTCAT
TTCCATATGTCTACCTCCTAAAAAATAAAAAGAACATCCATCCTATAAAAA
ATAGGACGGATGTTCTGTTCCGTGGTACCACCTATATTCAAGAAGAAATAA
TTCAATCTTTTCTTCAAGCACTTAAGTCTCTGATTAACGCTCATACACGCT
TCTACCTACTTTGTGCTCTATGTTTCAATAGAAGTTATAAAATAGGGCTACC
TTCAGACTTAATCATTTACATATTCACAGCCACCATATGCTCTCTTTAAAA
40 TGATATAAACTTACTCTTCCCAAATAATATATCTTTTACTCATTATATTT
ATAATGGATATTAATAAGTCAATGTATTGGAGATGAAATTATGAATGAA
TGTGCTTTTAACACAACGTATCCTATCTATATAGAGTATCATGATTATTAT
TGGGGACAACCAATCTATGATAGCAAAGAATTATTTAAGCTAATGGCTCTT
GAATCACAAACATGCAGGATTATCATGGTTGACAATATTAAAGAAAAAAGAA
45 TCGTATGAACAGGCATTTTACAACCTTTGAACCTCAATTTATCGCACATATG
ACTGAACAGGATATTGATTATTTAATGAAATTCCTCAACATTATTCATAAT
CGTAAAAAGTTAGAAGCAATTGTGAGCCAAGCGAAGGGATATTTAAAAATT
GAAAAAGACTATGGTAGTTTATAGTAAATTTTATGGTCTTATGTAAATCAT
CAACCCATAAATATGGGTTATAAAAAACCTAGAGATCGTAAAAAGTTGAT
50 CAAAGAGCTACTCAATTATCAAAAGATTTAAAGCATATGGTTTAAATTT
TTAGGTCAGTAAACCGTGTCTCATTCTTAGAAGCTGCTGGATTGTATGAC
TCACATCTTGAAGGATGTCTTTCAAACCAAAATCATGAGTGAAGCATAGAC
ATAAGTTGCGCTGCGGATATTCAAACAATATGAGTTTTTACGCATACTAA
AAGTTTACAACGTTTAAATTTATCTAGATTTTACTTTTAAAGTAATTACAATT
55 ACTCGCAAAATATTTTCGAAGTTTATTTAAAAATGTATAACTTAAGTCATT
AAAATCATTTATTATTTAATTTAAACATTAATAAAATAACCAATTAGTGGCG
CAATGACAAATAATATAAAAAATATTCTAAAAATGTGATAGCTTGAAATCA
TCGCTACATCAGCTCCAGTAGCCATAGCCACTAAAACATTGACTCATAC
CTCCTGGTGTCTCCTAAAAACAATTCATTGATGGATTTCATTAGTAATCA

AATGTATTCTATTATCATTATAAACGTTGTGACTATGAGCATTATATTTT
 GAAAGGCTATTGCTATTGCAATTCCTCCCTTTAAATCACTCATTAAGTTGG
 CAATCTGTAATCCAATACGTATCATATAAATAAGTTGCGCTGTGGCTAACA
 ACCAATGATCTAGTGAAAATGTAAATGTGTGTCATATTCCATATAATTA
 5 AAACATATTATAGGTGCTAATAATTGTTTAGTGGGGAAGTTAATTTTGACA
 TTCCATATATAGATGATTCCCACCATTGAGAATAAGATGATTATTTGCCATA
 TATTTAAAGTCTGAGAAAGTGTGGGTACTTCCATTGTAGTATGATTCATTT
 CATGATGGTTATCCTGAAAAAATACGAAATAAGTGGTACTAAAAATAACAA
 CAAATATTACACGTGATGTCTGTGTAAACTCACAACTAATATATTTGCTT
 10 TCTTATTTTCTCTGCCATCATAACATTTGGCTTAGCGCACCTGGTATAA
 CAATTAATAATCTACTTTCTAAATTTACTTGTGCAATTTTCTTAAAAAAGA
 ATGCAATTATCAAGCTAATAAAATTAGTAGGATAGTGACAAAAACGATAG
 TTAGCCAATTTTACTTTATGTCTTAAATCACTTGTGTGTGAAGGTAGAGC
 CAATTTGAACTCCAAGTAGTATTAAACCAATTTGACTCAACCAAAATGGCC
 15 ATCGTATTTTAAATTTTAACTTTTACACATAATAACGCCGCTAATATTG
 GTCCAAACATAAATGGCAAAATCACATGTAACATTTTAAATAACAGACTTA
 ACATAATAGCAATTAATAATATGATTAAATTATTAATAAAATTCGCTTCA
 AAATACCACCTACTTTCTGTTAAAAAGTAAGTCACAAAATTTCTAACTTTAT
 TTAAATTCAAAGAGTAGAACAAAATGCATCTTAAATTCCTAGTCCACCATC
 20 GAACCCCAACTTGCCCTGTCTGTGAATTTCTCAGTGAAATTCCTTTTGTT
 AGGGCCCCCAACCCCAACTCGCTTGTCTGTTGAATTTCTAAGTGGAATTT
 ACTGTATTGATGACATAACATAGCTTTAAAGTTGATATAAGTGCATTTT
 AATTCAGTTATAAACTACTTATAAGTTAAGAAAGCTAGGACATTGACTTT
 TGCCCTAGACTCAATTAATTTAAATAATTATATTTTAAATTTATCACACAA
 25 TACGACTTAATGCATTATCAAATGCTTGGATAGTTTCTCAATATCATCTT
 TAGTATGTGCAGTTGATAAAAAATGTTCTTCAAATGTGAAGGTGGTAGGA
 AAACACCTTCCTTAGCCATTTCTCTATACATATTACTAAATAATTTTAAAT
 CACTTTTATTTGCTTCCCTCAAATTTGTTACAGGCCCTCATTTAAGAAAGT
 AACCAATCATTGAACCAGCGCGATTTACTGTGATTGGAACATTATGCTTAG
 30 CAAATACCTCTTTTAACTCTTTTCAAGTATATCTCCTAGAGAATTAAAT
 ACTCATAAGATTCAGGAGTAAGTTGACTCAATGTTTCAATAACCACTAGTCA
 TTGCTAAAGGATTACCTGAAAGTGTGCCAGCTTGATAAATAGTCCCAACAG
 GAGCAATGTAATCCATAATTTCTTTTTTACCACCAAAAGCTCCAACGGGT
 AACCTCCACCTATCACTTTTCTTAAGCAAGTTAAATCAGGTGTTACACCAA
 35 AGTATCCTTTCGCGACAATTATAACCTACACGGAACCAAGTCATCACTTCAT
 CAAATATAAGTAATGCTCCATATTCATTAGTAATATCTCTTAAACCTTGTA
 GAAATCCATTCACTGGAGGCACTACACCCATATTTCCAGCAACCGGTTCAA
 CAATAACACCAGCAATATCATCGCCATATTTTTCGAACGCTAATTTAAGTG
 AATCTAAATCATTATATGGCACCGTGATAGTGTTTTAGCAATACCTTCAG
 40 GGACGCCCTGGTGAATCAGGTAAACCTAGTGTTGCAACACCTGATCCTGCTT
 TAATCAATAAAGAATCACTGTGTCCATGATAACACCCCTTCAAATTTTATAA
 TTTTATTACGTCTGTATAACCCCTAGCTAAACGAAGTGTGTCTAAAGTAG
 CTTTCAGTTCTTGAGGAAACCATTCCTTACTTTTTCAATTGAAGGTACACGGT
 CAATCACAAAGTTCAGCAAGTTTATTTTCTTGAAGTGTGAAGCGCCGAAGC
 45 TTGTACCTTTATCTACTGCTTCATGTAATTTGGATATAACTTGTGATTTT
 TATGTCCCAGAAATTAATGGGCCCCAAGTTAGCACATAATCAATGTATTCAT
 TTCCATCAATATCATATATTTTAGATCCTTCACCATGATCCATAAAATAG
 CTGGTGTGTCTACTGATTTAAATGCTCTTACGGGACTGTTAACACCGCCAG
 GCATTAATTTCTCAGCTTGCTCCATTGCTTTAATAGATTTTCAAACCTCA
 50 TCATATGACCTCCTCAATTATCATCTATTTATCTAAATAACGACAGATATC
 TTTTGCAAAATAAGTAATTTAAATCAGCACCTGCACGTTTCATAGATAT
 CATTTGTTCCATAACAATTTCTCTTCATCTATCCAACCATTTAACGCTGC
 TGCTTTTGTGCATACTATATCTCCACTAACGTTG

55

Sequence 3401

step.1002e12.cons.ok

GGTAGGTGTATTATTTATATTTACTACTTTGAATATGGCTTCAATTTCTGA
 TAGAAAAGTACATGCCAAAAATAAAAATATAAGAAGATAGGTGAGATAAAG

TGGAAACAATAGGAGTATCATATATTTAAAAGAAGGTTCCCAAAAGTTAAT
GATTATCAATCGAGGACCTATGTAGATATAGATAATCAAAAATATATATT
TGATTATTCAGCATGTAATTATCCTGTCGGTGTTCGTAGAAAGATCAAAATATA
CTATTTTAAACGAAGACAATATCGATAAAAGTCGTTTGAAGGTTATTCTGA
5 TCAAGATGAGATGAGATTTCAGAATTATTTAAAGAAATGAAAAATAATTT
AGATGATGATATTCAACAAGGTATTGTTCAAAAGCAAGATAATTTAGGTTT
AATTTAACATAAGTCATGATTATCTGAAGTTATCTATTATCGAATATTCAG
AGATGATATAATATATGAGTCGCCTATCTCTCAGGCGTCAATTGACGAAGA
GAGGAGGTGCATTAATTGCTAATCATTTTCGTTACATCACGACCACAGTC
10 CTCAGTTGTTGATATTATGTCATTATTTACTCATTTGGATATGTAATTGCAAT
AATAAATAGGCAACCATAGACACACATAAAAAATCCCCACACTATTGCAGTA
GTGTGGGGATTGGTGCATTAAATGCTAATCATTTTCATTATCTAGATTATA
ACACTGGTTAGTAGAGAAATGCAAGATAGACAAATAATTAATGACTGAGG
ATAAAGATATAAAACAAACATACTTTCTCATTTTGATACCATAGCCACCA
15 ACTTTTTACTTGTGTTCTATTTCATATTTAAATTCCTTCATAATTTCTTT
TGTTGAAAATTTCTGCTGCTTTTCATTTCAACGGCTTTATACTTTGTTCCTAC
TGAATAAGAAGCTCTTTATTTTGGAGCAACAGGTGTTTTAAGCTATCGAAA
ATTTTCCCATTTATTTATAAATTTGTTTCAAAAGGATACTTTGAAAATTC
ACTAACAAATTTATGGCAGTGGTTTAAATGAGTTAACTACTGATGAATTTAG
20 AGAAAAAGTCAAAAATGCTATTCAAAAATAATATTGAAAACCAAGGAAAT
TGGTGCCTTTTAAAACGTTTATTTTATAAACAACAGGACGCTAATAGTAA
AGATAGCTACCAAAAATTTTGGAAATGAGTTTGAAGTACTAGATGATAAGTT
TGACCTAAAAGAAAAATCGATTGTTCTATCTTGCAATGTCCCTAAAATTTT
TGGAGTTGCAACAAACCACCTGAAAGAGTCTGGATTACGAACGTAAAAGG
25 GGTGATGCGAATAATTATAGAAAAACCATTTGGTGATGATTTAAAATCTGC
AAAAAATTAAATAATCAAAATAAGAAAGTCTTTTAAAGAAAAAATATTTG
GAATTGATCACTATTTAGGTAAAGAAATGATTCAAAATATTGAACGTCTAC
GATTTTGAACACTATATTTGAACCACTTTGGAATAATAAGTATATTTCCA
ATATACAAACCACTTCTTCAGAAACGATTAGTATAGAATATCGTGGTGGTT
30 ACTACGAATCAAGCGGCGCCCTAAAATTCATGGTTCAAAATCATTTATTAC
AAATGGTTTAATTACTTGCAATAGAACCTCCAATTAGTAGAAAAAGTAGTG
ATATAAGAAAAAGAAACATTTCAAGTCTTAAAATCCTTGAAATGTTTAAAT
CCAAATGAAATTAAAGAAAGTTTCGTTTCGTGGTCAATATGATGGGGGAATG
ATGAATAATGAGTTTCGTTCTGCATATAGAAATGAACCTAATGTAAATTCA
35 CAATCTAATACTGAAACTTTTGTAGCAGGTAAAATAGAAATTTGAAAACCTCT
AAATGGGCTAGTGTTACATTTTATATTCGTACAGAAAAAGAAATGAAAAAA
ATCTATCCAAATCGTTATAGAGTTTAAAAAATCCAATGCAATTTTATTATA
GTAATTGTGAACAGAAATGCATCAAACTTATTAGTCATCAACGTACAACCTA
ATGAAGGATTTTCTTTATGTGTGAATGGTAAGAAAAGTAATCAAAATAATG
40 AAATGCAAAAAGTGAAGCTTTCTTATACTATGCCGATTAAAGATAAAATGA
ACACAGTTGATGCATATGAAAATCTTATTTACGATACATTAATTGGAGAAC
AAACAAAATTTACGCATTGGGAAGAATTAAAAATCTTGGAATTTATTGA
TGATATTGAAAATGTATGGAAACAAGAATAGCCACAGTTTCTTAATTATGC
CTTTGGATGCTATGGGCCATAAGAAAGTGA AAAATTA CT TAGTGAAGACGG
45 ATTGAATTGGTGGAATAAATTTGTAATATAAAAGACCTAAAGTTGTCGTGTT
TTGAATGAGTAGTTAATCCCTTAATTAATATGCAATAATCAATGTTATAAAG
TTTATAAAGGGAGGACAATTAATGGATTTTATTAATATTACAGGTGCTTCA
CAAAATAACTTGAAAAACATAGATGTAAATATCCCAAAACACTTAGTAACG
GTATTTACAGGTCTGTTCTGGTTTCAGGGAATCATCTTTAGTGTTAATACT
50 GTTGCTGCGGAGTCTGAACAGCTACTAAATGAAAGTTATTCTAGTTATATT
CAATTTTCATTTAAATCAACAACCCAGACCGAAAGTAAAGAAAAATAAAAAT
CTTCTCTGATGATGACGATTAAATCAGAAAAGATTCAATGGGAATTTCTCGC
TCCACGCTAGGAACAGTTTCAGATATATATGCTTCTGTTAGATTACTGTGG
TCTAGAATAGGCGAACCGTTTGTGGTTATTCAGATGCATATTCCTTCAAT
55 AGTCCTAAGGGCATGTGTAAAACCTGTGAGGGATTAGGATATATTGAAGAC
ATTAACCTTAGATGAATTGCTAGATTGGGATAAGTCTTTAAATGAAGGTGCA
ATAGACTTTCTTCTTTTGGACCAGACAAAGAGCGTGGTAAAGCCTATCGA
GATAGT

Sequence 3402

step.1002f01.cons.ok

5 GTATTCAACAAGTTTTTTTAAATTTATCTTTTCGTGTTAAATGTCTAAATCC
TTTATCTAAACTTTTTCATTTATGTACTCCTTTAAAAATATCTTTATATTACT
GTAGTAATGTAACCTTTTAACTTTTCGCTTTAAATATCAACTGCTAACATTG
TCTTGCACCTCAAATAATTCAAAACCTTTTCACAATTTTGTAAAGCGCTTAC
ATATTTATTATATATTCCTGGTCATGTCTGTTATATTTTCATTGTATAAAAC
10 GAATTTCTTTGGTATGATACTAGTGCTACTTCAAAAAATGAAAGGATGTC
ACACCTTATGAATATAGGTATAGATAAAATAAGTTTCTATGTACCCAAATA
TTATGTAGACATGGCTAAACTTGCAGAAGCGCGCCAAGTCGATCCTAATAA
ATTTTAAATTGGAATTGGTCAAACCTGAAATGACTGTGAGCCCAAGTGAATCA
AGATATCGTATCTATGGGAGCCAATGCTGCTAAAGATATTATAACAGAAGA
AGATAAAAAGAATATTGGTATGGTTATAGTAGCAACTGAGTCTGCGATTGA
15 TAATGCCAAAGCAGCAGCCGTTCAAATTCACCATCTTTTAGGTATTCAACC
CTTTGCAAGATGCTTTGAAATGAAAGAGGCTTGTATGCAGCAACACCTGC
AATTCAACTTGCCAAAGATTATCTTGCTCAACGCCCTAACGAAAAGGTTCT
TGTCATTGCTAGTGACACAGCTCGTTATGGTATTCATTCTGGTGGTGAGCC
TACTCAAGGTGCCGGTGCAAGTTGCAATGATGATTTACATAACCCAAGTAT
20 TTTAAACTTAATGATGATGCCGTAGCATATACTGAAGACGTTTATGATTT
CTGGCGTCCAACGGGTCATCAATATCCCTTAGTTGCTGGTGCATTGTGCGAA
AGATGCCTATATCAAGTCATTCCAAGAAAGTTGGAATGAATATGCACGTCG
CCATAATAAAACACTCGCTGATTTTCGCTTCACTATGTTTCCATGTACCATT
CACCAAAATGGGACAAAAGCTTTAGATTCATTATTAAATCATGCCGATGA
25 AACTACACAAGACCGTCTTAACCTAGTTACCAAGATGCAGTTGATTATAA
TCGTTATGTCGGTAATATTACACAGGGTCCCTATATTTAAGTCTCATCTC
TTTATTAGAAACACGTGATTTAAAGGGCGGACAAACGATTGGTCTCTTTAG
TTATGGTCTGTTCTGTAGGCGAGTTCTTTAGTGGAACATTAGTAGATGG
ATTCAAGGAGCAATTAGATGTTGAGCGCCACAAATTTTATTAAATAATAG
30 AATAGAGGTTTCTGTTGATGAATATGAACATTTCTTCAAACGCTTTGACCA
ATTAGAATTGAATCATGAACCTGAAAAATCAAATGCAGATCGTGACATTTT
CTATTTAAAAATCTATTGATAACAATATTTCGTGAATATCATATAGCAGAATA
ATCTAAACTCAATTCCTTTTGCTTTTTTAATTTTTTAAATGAAAGAAACGA
ATTTACTATTCTACTTTGCCCCCTAAAAACAACGTCATATAGCTGTTTTTA
35 GGGGGTTTTTAATCCCAAGATAGTGCATTTTCATCTATTATAGACCTCTACT
ATTTATATACTGTTTACCTACTATTTCATAGATAAAAAACGAATCCTATCCAT
AGATCAACATTGACAAATAACTAAAGCAATTCATACAATAACGTATCTCA
AGGTTTCGTTGTTTTTTTAGGTCTGTATAAATGTGTCATATCTATTCTTC
CAATTTTAATAATTTTTTATTTGTTGTCTAAAGTACCTCCGTATCCTGTTAA
40 ATTTCCATCTTTACCAACCACTCGATGACATGGTATAAGTATAGAAATAGG
ATTACTCCCAACAGCACCTCCTACAGCCTGAGCAGACATTGCTGGCTTATT
CATTGCTCGTCTACTGTTTATAGCAATCTCTCCATATGTTTAAAGTTCTCT
ATAACGAATTGCTCTGAAGTTCATTCCATACCTTCTGTTGAAAAATCAGTGCC
TTTTGGTGCTAACGGTACTTTTATCTCAGGATAATTCCCTTTAAAAATACTC
45 TCTTAACCACGCTTTAGTATCTTTAAAAACATCTAACGTATCCTTTGTTTT
AATTCCTTTCAAATGTTGATTAGGGTATAAGACATGTGACAACGAGATGCC
ATCGTTTGTAATTAATTGCAATAAGCCAACGGGCGAGTCATACATTGATTG
ATACATATTTCTCCAACCTCCTTAAATGTGTATATCATTCTAGTCATATCCT
TATAAACTAACGCATTCTATCTGAACATAATTTTATAATATAGATAATTG
50 CTATATTTATGAAATTGGATAAACATTCGAATAGAATGTGATACACCAAAG
AATTGTACGTTTTATTATCTTTAGCAACCTCTTTTAAATTTCTTAGTATCT
TCATGCTTAACATAATTATTTAATTGCTTTTCACTTGAATGACACGCCCAT
GAATCAAAATAGAAAAAGTTTACAAAAATATAACAAATTGCTAAAAAAATT
ATTATACCAATAACATAGAATACTTTTTTCATTAAAGTTACCTAGTTTTTC
55 AAAAAAATTTTTATAATTTATGATAACTTAAAGGTGTATATTTTTCTACA
TAATATATTTAATTTTCAAGGAGATGTAAAAAGTTGAAAAATTTTCGCAAA
ACTAATTTTGGTGGGCATTTTAGTATCGGGGTGAGGATAGCGAGTGATACA
AACAAATATAACTCACGCAAAAGAAAGTCACGATTCAACTCCTCAAAATAT
TAAATTAGTGGGAACGTATGATACTTCTCAAGTTGATTCCAAAACGATGAA

ACAATTTAAAGAAATAGAAAAAGAAGATAATAATTTCCACAT

Sequence 3403

5 step.1002f02.cons.ok
 CAGAAGGATCAATCACACCATAGCGATGTGTATCAGATTCTGGCACTTCTT
 GAACACCTATTACTGAATGGCCGTGTTCTTCATAAACATCCATTAATTGTT
 TAATAGCTGGTGTATCAGACTCTACAATGTATCTCCTAATAACACTGCAA
 ATGGTTCGTTACCTATAAACTGTTTTGCAGTATGAATTGCATGTCCTAGCC
 10 CTTTTTGTCTTTTGTGCGACGTAAAAAATATTAGCTAAATCTGTTGAAT
 ATTGTACTTTTCAAGTAAATCTGCTTTTCTTTACTTTCAAGTACTATCT
 CTAGTCTTTTGTGATTGTCAAAGTGATCCTCAATTGCACGTTTATGCTG
 CAGTCACTATAATAATATCTTCTATTCCTGCATTAAAAGCTTCTTCTACAA
 TATATTGAATTGTTGGTTTATCTAATATTGGTAACATTTCTTTTGGCATCG
 15 CCTTAGTTGCTGGTAAAAACGAGTACCTAAACCAGCAGCTGGTATGATTG
 CTTTTTTTATGTTTTTCAACGTATAAACTCCTTTAACAACCTAAGTTTTTG
 ACTACTTAATAAAGTGTATCATACTATTAAAAATGCTTTCATAGTAATGTTT
 TTCACCACAATCTTTTTCAATATTAACATTAACCTAATCAAGAATTGAGTT
 AGAAATTAATAAATACTCGAATAGAAATTAACCTTGTTTTAAATATATCGGT
 20 GACAGACATCTACTGAAAAACACTCGCTATATTTTAAAAATGAGAGCTTCA
 GGAATTTTCTCTATACAAATAAACTGATAAATAAAATAAAATAAAATACC
 AGATACTGAAATAATTGCTCTTAATAACGACCATGTTAAAAATGTTTCTTT
 AACTGTGACGCGGAAATATTCCTTAAACATCCAGAATCCAGCATCATTGAC
 GTGAGAGAGAATTACGCTACCTGCACCTATTGCAAGAACGACCAAAGCTAC
 25 ATTAACATCTGAATGTTCTTAAAAGTGGTAACACAATGCCTGTTGTGATAC
 GGCAGCAACTGTAGCTGATCCTAATGAAATCCTTAAGACTGCAGCTACAAT
 CCATGCTAATAAAATGGGCGACATGCTTGTTCCTTCAAACATCTTAGCGAT
 TGTATCACCTACGCCACCATCGATGAGCACTTGTTTAAATGTACCACCACC
 ACCGATGATGAGTAACATCATGCCGATTGGATAAATAGCATGCGTAACTGA
 30 TTTTCATAATGTCTTCCATTTTCCGTGTTGCTTCATTCCCATTGTAAAGAT
 TGCAAAATATCACTGCGATAAGCATTGCAAGTACCAGCAGTTCTTATGAAGTA
 AACAATTTGTTTGAATACATTTGTTGCTTCTTCGTGTCCTGTACTATGGAT
 GGTGACGGTGCAGCTGCAATGCCGTTATACCGAAGCACGTATGACTAAAAATA
 ACATTAGAACTTTTACGTGATATTAACAAAGACACAATTGATTTTATGAC
 35 AACTATGATGGTAATGAAAGAGACCCGTACGTCTTACCTGCACGTTTCCCT
 AACTTACTAGTAAATGGTGCAGGAGGAATTGCCGTAGGTATGGCTACAAAT
 ATTCCTCCCCACAATTTAAGTAACTGAAAGTTATTGATGGTGTGCTCAGTTTAAAGT
 AAGAATCCAGACATACAATTAATGAGCTGATGGAAGACATACAAGGTCCT
 GATTTTCCCTACAGCTGGTTTAGTACTAGGGAAAAGTGGTATTCGTCGAGCT
 40 TATGAACAGGTCGTGGGTCAATTCAAATGCGTTCTCGTGCTGAAATAGAA
 GAACGTGGTGGTGGCCGTCAACGTATTGTCGTAAACGGAATACCTTTCCAA
 GTCAATAAAGCGCGTATGATTGAAAAAATCGCAGAGTTAGTTAGAGATAAG
 AAAATCGACGGTATTACAGATTTACGTGATGAAACAAGTTTGGCTACAGGT
 GTAAGAGTAGTTATTGATGTACGTAAAGATGCAAAATGCGAGTGTATTTTA
 45 AATAATTTATATAAAACAAACGCCATTACAAACATCATTTGGTGTGAATATG
 ATTGCTTTAGTGAATGGTAGACCTAAACTAATCAATTTAAAAGAAGCACTT
 ATCCATTACTTAGAACACCAAAAACAGTGGTTAGACGACGTACTGAATAT
 AATCTTAAAAAAGCAAGAGACCGTGCCCATATTCTAGAAGGTTTACGAATA
 GCACCTAGATCATATTGATGAAATATCACAACAATTCGTGAATCGGACACT
 50 GATAAAATTCGATGGCAAGTTTACAAGAGCGTTTAAACTAAGTGAACGT
 CAAGCTCAAGCAATTTTAGATATGCGTTTAAAGACGTTTAACTGGATTAGAA
 AGAGATAAAATAGAATCTGAGTATAATGAACCTTCTAGAATATATTAAAGAG
 TTAGAAGAGATTTTAGCTGATGAAGAAGTACTATTACAATTAGTTTCGTGAT
 GAATTGACTGAAATTAAGAAGCTTTTCGGCGATGAACGTGCGACTGAAATT
 55 CAATTAGGTGGTCTAGAAGATCTTGAAGATGAAGACTTAATCCCTGAAGAA
 CAAATTGTTATTACATTAAGTCATAATAACTATATTAAACGTTTACCAGTA
 TCTACATATCGTTCTCAAAATCGTGGTGGTGGTGGCATAACAAGGTATGAAC
 ACGTTGGATGAGGACTTCGTTAGTCAATTGGTAACAATGAGTACACATGAT
 TATGTTCTGTTCTTTACGAATAAAGGTGCGTGTATATAAACTCAAAGGTTAT

GAAGTTCCTGAGTTGTCACGTCAATCCAAAGGCATACCTATTATTAATGCG
 ATTGAAC TCGAAAATGACGAAACAATAAGTACGATGATTGCAGTTAAAGAC
 CTTGAAAGTGAAGAAGATTATCTCGTATTTGCGACAAAACAAGGTATCGTT
 AAACGTTTCATCATTAAGTAACTTCTCCCGTATTAACAAAAACGGTAAAATT
 5 GCAATTAAC TTTAAAGAAGATGATGAATTAATTGCAG

Sequence 3404

step.1002f03.cons.ok

10 GGACAGTATTCCTATTATCACACCCATAATAATGGGACGAATACTGGTTTG
 AACTTGTAATAATGAGATATTCATCATTATTAACGAAACACATATCACTGT
 AAACATATAAACTCAAAAGTTAGAGTTAAATGATTACTTTTAAAGAATAC
 AACATTTCATAAGTGAAGCAGATTAATTAATCCTATACTAGCTGCTGTAAT
 TATTACTACCGTGATTTTAAATGATGCATTTGCATAACGATTCATATGAAT
 15 TTGATTATGTTTCACGAATTGCTTGAGTAAGTAATGGGATAAGAACGAACT
 AAAAGTCGTAGTTACAATCAAACCCATTTGTATAAATGAAGCACCACGATC
 ATAAATGCCCTTTTGAATAATTGCTTCTTTAAAAGCAATACCGCTATGTTG
 TAATAAACGTATTATTGTAAAACATCCACAACCTTGCCATAAAAATAACGAT
 AAGTTGACTCAATGCAATATGGATATGGAAATAAACAACCTGCTTCCATTG
 20 AATGGAAGTATTATTAAAGCGATAGCATAACTTAAGTTTAAGTGGTTT
 AAGTAATAAATATAACATTGAACCTAAAAACCAATCGAAGATGCCAATAT
 AGCTAATGCTCCTGCTTGATAAATAGACCAGTGTTTCATTGAAAACATAAT
 AATTGCAACAATGATTAACTAACTCTAATTACCTGTTCTATAACCTGGGA
 AATTGCTGGTATGGTCATTACTTGTGTTTGATTGATAAAAATCCTCTTAACAC
 25 TCCTAAGACACCTATTAAAAATAAACTAAACCTGGCCATCTTTAACATGGG
 TGCTAAATTAGAATCGCCCATCCATCGGGTAATCATATTTCGCAAACATAAA
 AAGCAAAATAAAGACGATAAAACCTATGCATTGTAATCGAAACATAACCTT
 TGTATAGACTTTCATCGGATCGATTAAACACCTATCACTTGAGTCACAGCACT
 TGGAAATAGCATTCATAGATAAAATAACCCCTAGTGCTACGACAGGATATAT
 30 TTGTTGATAAGCATATAAACCGTCATCACCTAAAACATTTTGATACGGAAT
 TCGATAAATGGCACTTAGTATCTTCACAATAATTAATGCTAAGGTAAAGTAT
 AACGACGCCGTAAACGCCGTCTCACTCTTAGTCTTCATCTTTAATCGCCA
 TACTTTCTTCAATACATCTAACTAAGAATTTCAAACATCTAACCATTGTT
 TTGATTTTGTAGCGTTACATTCATTGCATTATTTTGCACGCCAACCTTCA
 35 TTGCTCTACCAAGAGGTTGCGTCTGTTTAAACAATCTCTTCCATTAAATAT
 CTTTCAGTCGCTTTAGGTGATAAATGATTGTTGATAGATTGCGCTTTGTCTT
 TTATCAATTTCGACACCCGCATGTAGAGCGTGGACTTTGATTTCACAATAT
 CTAATAATCGTTTCGACTTCAATTGGATAATCATTAACGATCTATTAAAT
 CATCTTTGACATCGAAAAGTTGTTCTTCAGTTTCTACTTTTCGAAGTTT
 40 TATAAATCTCAATTTTAGCCTGTTCACTTTGTATATATTTCAGCTGGTAAAT
 AAGCATCTAAGTGCAATTCATTCAATATCTGGTGATCCGGCGATTCTT
 CTTTAATGCCACGTTTTTCGTTTACTGCTTCTTCAACATTTGAGAGTATA
 AATCGAAACCAACCGAATCAATAAAGCCATGTTGTTGCTTACCGAGTAAAT
 TGCCCTGCACCACGAATATTAAATCTCGCATAGCGATTTTAAACCTGAAC
 45 CTAGTTCCGGTAAACTCCTTAATAGCTTGCAATCGCTCTTCAGCAGTCTCAT
 TTAACACTTTGTTAGCTGGATGTAAGAAATAAGCGTAACCAATTCCTACTTG
 AACGTCCTACACGTCCTCTTAATTGGTATAGCTGGCTTAAACCAAAACGAT
 CAGCCCTCTCTATGATTAAAGTATTAGCATTTGGTACATCTACACCTGTTT
 CAATAATTGTAGTCGTTACTAAAATATCGTACTCGTGATTAATAAAGCTTA
 50 ACATTGTTTCCCTCAAAATCAGTTTCAGTCATTTGGCCATGTGCTACAGCAA
 TTTAGCGTCAGGCATTAAACCTTTGAAGTTGTTCTTTTTTCATAAATGG
 ACTGCACCTTGTATACAAAATAAAATACTTGTCCATCGCGAGATAATTCAC
 GCTCTAATGCCTCTTTAATAAAGTTTCGTATTCTGTTCTAAGACATAAGTTT
 GTACAGGAAAACGATTTTTCAGGTGGTGTTCATCACTGATAAGTCACGTA
 55 CACCTAACATACTCATATGCAATGTTCTTGGTATTGGTGTTCAGTAAGCG
 TCAGTACATCAACGTTTTTTTTCAAAGTTTAAATGCGTTCTTTATGTCGCA
 CTCCAAACGTTGTTCTTCATCAACAATAAGCAATCCCAAATCTTTATATT
 GAATATCTTTACCTAATAATTATGTGTACCTACGACAATGTCAACATATC
 CTGATTTGAGCCCTTCTTTAGTTTCCCTTATTTCTTTAGCTGTGCGGAATC

5 GACTTACCAATTGAATTTTCGACCGGAAAACTTGCATACGTTCAAGTAGTG
TTTCGTAGTGTGTTGTGTGCAAGAATCGTTGTTGGTACTAAAAACGCAACTT
GCTTACCATCCATGACAGCTTTAAACGCTGCACGGACAGCCACTTCAGTTT
TACCATATCCCACATCACCACAAAGAAGACGGTCCATAGGTCTTGCTCTTT
10 CCATGTCACCTTTAATTTTCATCGAATGACTTACTTTGATCGGGTGTAAAT
CATATGGAAAAATCATGCTCAAATGCTGATTGCTCTGCAGTATCTTGGCCGT
ATTGATAACCTACAGACATTTACGTTCTTTATACAAATCAATTAATTCAT
CAGCTATATCTTCAACACTTTGTTGAACTTTTCGCCCTTTGTTTTTTTCCATT
CTGTCCCGCCTAATTTATTTAATCTAGGCGACTTATCTTCAGAAGCCACAT
15 ACTTTTGAACCTTGATCCATTTGATCAACTGGAACGAATAATTGATCAGTTC
CTTCCACCTCGAACCATTAACTTCGTCATCTTCGAGGGATCTTATAACCG
AAGTTGGGAAATCTCATCTTGAGGGGGCTTCATGCTTAGATGCTTTCAGC
ACTTATCCCGTCCATACATAGCTACCCAGCTATGCCGTGGCAGCACAACCT
GGTACACCAGAGGTATGTCCATCCCGGTCTCTCGTACTAAGGACAGCTCC
20 TCTCAAATTTCTTACGCCCACGCGATAGGGACCGAACTGTCTCAGCAGC
TTCTGAACCCAGCTCGCGTACCGCTTTAATGGGCGAACAGCCCAACCCCTG
GGACCGACTACAGCCCCAGGATGCGATGAGCCGACATCGAGGTGCCAAACC
TCCCCGTGATGTGAACCTTTGGGGGAGATAAGCCTGTTATCCCCGGGGTA
GCTTTTATCCGTTGAGCGATGGCCCTTCCATGCGGAACACCGGATCACTA
25 AGTCCGCTTTTCGACCCTGCTCGACTTGTAGGTCTCGCAGTCAAGCTCCCT
TATGCCTTTACACTCTATGAATGATTTCCAACCATCTGAGGGAACCTTTG
AGCGCCTCCGTTACCTTTTAGGAGGCGACCGCCCCAGTCAAAC TGCCCGCC
TGACACTGTCCTCCACCACGATAAGTGGTGCGGGTTAG

25

Sequence 3405

step.1002f05.cons.ok

GTGGTATTAAACCAGATAC TATCAAAGATATTGTTGCTGAAGATCCAGATT
TAGTTATTGTTGGTGGCGGTATTGCGAATGCTGACGATCCTGTAGAAGCAG
30 CAAAACAATGTAGAGCAGCTATTGAAGGTAAATAAGATGAGTGAATTTAAT
AATTATCGTCTTATTCTTGAAGAGTTAGATTCTACTTTATCTCAAGTAGAT
AATACAGAGTATGAACGTTTTGCTAATGATGTTATAGGTGCAGATCGCATA
TTTACAGCTGGTAAAGGTCGTTTCAGGTTTTGTTGCTAATAGTTTTGCAATG
CGCTTAAATCAATTAGGTAAAAATGCCTACGTTGTAGGTGAGTCAACAACA
35 CCTTCAATTAAAGAACATGATTTGTTTATTATTATTTCAGGTTTCAGTTCT
ACAGAACATTTAAGATTATTAGCTGAAAAAGCACAATCTGAGGGTGCAAAA
ATTGCTTTATTAATACTACAAATGCGGAATCGCCAATCGGTAATCTTGACAG
ACGGTTGTGAATTGCCTGCAGGTACTAAACATGATGTTGAGGGTTGCAAAA
CAACCACTTGGTAGTTTATTGTAACAGGCTTCACCTATATTCTTAGATAGT
40 GTTGTATTACCTTTAATGGATGCATTTACATTAGTGAAAAACAATGCAA
GAGAATCATGCTAATTTAGAAATACTAGAATGAGAGATGAGCACTTTTGTC
TATGATTAGACAAAGGTGCTTTTTTAATAAATGGTGCAACATTACTCCTTT
AAGTGTATATTTATAATATAACTTTAACTATGAGGTGAACATCATGAATTT
TGATAGTTATATTTTGAATTTGATGGAACGCTAATTGATACAACAACATG
45 TCACGTCAAAGCTACGCAAAGCGCTTTTAAAAGATTAAATTTAGATGAACC
TACAGAACAAGCTATTTTACATACATATCATTTAAATTTATATAACAATTT
TAAAGCGTAGCTTCACATGAAGTGTCTTTTTATCAAATAGAAAAATTAAT
AGATGAATACAATCATTGTTTTCAGCAACGATGAAATACATCAATCAAAAGA
ATATACCGGAATAAGTGAAGCATTAAATTTTACATAACCAAAAGAAAAA
50 AATATTTGTAGTGTCTAATAAAGAAATACTAACAACCTAAAAGTATTTAGA
TTATCTCGGATTAAAGCCGTTTTATAACTGATTCATTAGGTGCTGTATTAA
AAATGAAGACAAACTTCTTTGTGAAACGATTCAAATTTGATACAGAAACA
TCATTTAATGATAGGTAAACCGTGTATATAGGGGACACAGCACAGAAATAT
CAAGAGTGCGAATCAAGCTCATGTGCAACATGCGCTGTCACATGGGGAGC
55 ACAATCTGCACACGAATTGTTGCATGAAAATCCTCATTATATTGTTAATGA
TCCAGAAGAATTTTTAACAATTTTATAATATTTTCAATGGTAGGATGATAA
TTCATGAGGAATTTTATCCTACTCTTTTTTTAAACTATGAAGCAAGTATA
ATGTAAACATAATTATATAAAGTGTTTACAAATGGTGATGAAGATGAAAAA
AATAATGGAATATTTACAGCATTATATTAATCAATACCCGCATCGATTAGC

TTTAGTGTGTTGAAGATCGACATCTAACGTATGGAGAATTAAGTAAAGAAAT
 TTATCAGGCTAGTATGCGCTATAAAGAAGTAAATTAACAAAAAAGTAGG
 TCTAATGGATGAACATCCTGTAAATAATATTATTAACATTTTTCGGGTACA
 TCAAAGAGGTGGAATTCCTTGCAATTTTAATCATCAATGGAGTAATGAAAG
 5 GATACATCAACTTGTAAGTTATGACATACAATGGTTAATTAAAGATAA
 TCATCTTACCTCAAATCATGATAACTCAATTTATAATGATGAGGTTATCCC
 ACGTAATGTTATACATATAGGTTTCACGTCAGGAACTACAGGTTTACCCAA
 AGCGTTTTATAGAAATGAACATTCTTGGATAGTTTCTTTTAAGGAAAATGA
 GAAATTACTCCAGCATTGTGAAGAAACCATTGTAGCACCGGTCCTTTATC
 10 ACATTCACTTTCATTGTACGCATGTATTTATGCATTAAGTACTGGAAAAAC
 ATTTATAGGTAACAAAAATTTTAATCCACTATCTCTTATGCGTCTTATTAA
 TCAATTGAACAAAACGACAGCAATATTTGTAGTGCCAACGATGGTACAACA
 ACTTATTTCAACTCAACGACATGTTCATCGATTAAAAGTATTTTGAGTAG
 TGGTGCTAAACTTACATTGCAACAGTTTCAACAAATCAGAAATTTATATCC
 15 ACAAGCAAATTTAATAGAATTTTGTGGGACATCTGAAGCAAGTTTATAAG
 CTACAATTTTAACCAATCATCTCCTGCTAATTCGTGGTAAACTTTTCCC
 TCATGTCGAGACGAGATTATTAAATCAAGATGATGATGCAGTAGGATTATT
 AGCCGTTAGAAGTGAAATGGTGTTTAGTGTTATGTTGGACAAAGCAATCA
 AGAGGGGGCATGGATTAAAACAGGCGACTTCGCTTATATTAATAATCAACA
 20 TTTGTTTTTAGTAGGTAGAGAGAGTGATCGTATTATAGTTGGGGGGATTAA
 TGTATATCCAACAGCTATTGAAAGCTTAATTATGGATATTGAAGGCATTGA
 TGAGGCCCTTGTCATTGGTATACCACATGCTAAATTTGGAGAAATAGCGAT
 ATTGCTTTATTCAGGTAAAGTACAATTGAATTACCGACAAATTAATCTTT
 TTTAATGAAACATCTTTCAAGACAAGAAGTTCCATCAAAATTAAGAAAAAT
 25 TGACCATATGATTTATACAGAATCAGGAAAGATTGCTAGAAAAGAGATGAA
 AAATAAATTTAATGGAGAGTTATAAAAAATGAAACAACCTGTTATTATT
 GCAGCAAAACGTATAGCTTTTCGGTAAGTATGGTGGCCGATTGAGGCATTTA
 GAACCTGAATCATTACTAGAACCTTTATTTAATCATTTTACAGATCAGTAT
 CCAAAGTAATGTCTCTTTTGGATGACGTCATTTTAGGTAATACGGTAGGT
 30 AATGGGGGAATTTAGCTAGAAAATCATTACTTGAAGCGGATTAGATTTT
 AAAATACCTGGTATAACAATTGATCGTCAATGTGGCTCAGGTCTTGAAGCC
 GTTATA

35 Sequence 3406

step.1002f06.cons.ok

TGAAGTTTAAGTCTTTTAAATAAATCTACAACAAAACTTTACGTTCTCCTT
 TGAAAAATTTACCGATTGCTGGGATAAATTTTTCATGATCTTTTAGCATAT
 ATATCAGGAAGAAAGGCACCATTTAATAAAAAATACAGTTGAGATGAAAG
 40 ACGTAATATATGACACTGAATTAGATAAAATAGACGTTGCGCCATCTCCCA
 TGGATTTAACCGCTTTATTAATACGATGTGTGACATCATCTGGTAATTTAT
 CCATTTGCTCTAACGAGAAATTAATTAATTGTTTCGGCTCTCTTTTGTAAATG
 ATGGTGTTTGTTAATTAAAGTTATTAATATTGGAAATAATGATAGGTGCTA
 TAAATGATACCACAATAGCGATGATAGCTATTAAATCCTATGAATATTGTTG
 45 TTATACTAGCCCCAACGTGGAAAGCCCCACTTTTCTAATATGTTTGGAAAAG
 GTAAACATATGTAAGAGAAATCCACTAATTAATAATGGAAGGAAGACTG
 AACCAATGATAGTAGCTATTGGAGCAAATACCTTCATGCACTTCCATAAATA
 GTTTGATGAGTATGAACAGCATAATAAAAAATATTCCTGTTCTAAACCAAA
 CCTTATTGAACATAGATTGTGATCCTCCTATTTTATTCTTTTTCATACTCA
 50 ATAAGTATATATCAAACTCAGTCATTAAAAAGAGTAAATACCTTTTACAAT
 ATAATGTCATGCACCTTTATTTTAAATAAAATGAAAAAAGTAGTATGAAAT
 AATTTGTATACGAGCTTGCTATAAAATTTAAAAGTATTATAAGTATTGATAC
 AATAGTATTAAACATTACAAGGGAGCATCGGTTTATGTGCGAGTTTAAAGTA
 GGTAAGATTAAATAAAAAAATAATACATAGTGATATTTTAAATAGAGATGTA
 55 ACATTATCCGTTTATTTACCTGAGGATTATACAACTTATTTAAATATCAG
 TTGATTTTTCGCTTTGATGGTTTAGATTTTTTTTAGATTGTTAGTAATACAA
 CGCATATATGAACAGTTACGTGAAGAACAATCAATACAACGTGCAATTATA
 GTAGGATTCCATTACGAAGATGTTGAAAAACGTAGGGAAGAATTTTCATCCT
 TCAGGAAGTCGTTCTAATTTAACCATTAAAGCAATGGGAAAAGAAATTCCTT

CCTTATATCGATGCGACATTTCCAACCTATAAAGTAGGTAATACAAGGTTA
 CTTATTGGAGATAGTTTTAGTCGAGTTGCGTTATTGAGCCCAATGTATAATGAA
 TACCCAACATTTTTAGTCGAGTTGCGTTATTGAGCCCAATGTATAATGAA
 AATATTAAGAAAAAATTGATACATGTATGAATAAAGGTCAATTGACGATA
 5 TGGCATGCCATTGGTTTAGAAGAAGCAGATTTTATTTTACCAACTAATGGT
 AAAAGAGCTAACTTTTTAACACCTAACCGTGAATTAAATCAACTGATTAAA
 GAAGATAATATTGAATATTTCTATAAAGAATTTAACGGTGGACATCATTGG
 AAATCATGGAACCATTGCTAGGAGATATTCTCTTACAATTTTTAGGTGAT
 CCAATAAATGGAAAATATGTTTAAATAAAGTAAGATAAATAAATAGTATGT
 10 TAGCGAATGGTTGTTGAATGCGAAACACTATTAATTTTCAGAAATGTAATTG
 TTTTCATGATAAAAAGTAACGGTTTTTCAAAAAGTTTGTATATAATAGGAATA
 AGTTAAACAAAGGAGGAATTTAAATGATTTTAGGATTAGCATTTGGTTCCGT
 CAAAGTCATTTCAAGATGAGGTGAATGCTTATCGCAAGCGATATGACAATC
 ATTATGCTCAAAATATGCCTCATATCACGATTAAACCTCAATTTGAAATCG
 15 ATGATCATGATTTTTAATTTAATTAATAAAGTAAGTGAATAATCGAATTTCTA
 GTATTAACCAGTAGAAGTACATGCTACAAAGGCATCTAATTTTCGCTCCAA
 TCAGTAATGTTATATACTTCAAAGTTGCTAAAACAGAGTCATTAGATCAAT
 TATTTAATCAATTTAATACAGAAGATTTTTACGGTACAGCTGAACATCCTT
 TTGTACCACATTTTACAATTGCCCAAGGCTTAACAAGTCAAGAATTTGAAG
 20 ATATATATGGTCAAGTAAATTAGCAGGGGTAGACCATAGAGAAATAATTG
 AAGAACTATCGTTACTTCAATATAGTGAAGAAGAGGACAAATGGACTATTA
 TTGAAACTTTTACATTAGGATAAAAAGTTCAAATTTGTAAAGTGAATTTGA
 ATTTACAGAATCATTATTGTTAAATACGTGAAGAGCGCTTAATCAACTAAA
 AACGCCAAATCCTATTGTGTTTCAGTGGGATTGGCGTTTTTATATGAAAC
 25 ATAATTATATTGTATAGTTAATCATTAATAAAGATTAATGTACGTTTTATA
 TTGAATATTCATATTTCAAATGCATATGAATGACGTTTCATTATATATAC
 AACTAATCTTTATATAAACGATTTATAGATTTTAAGGTTTAACCTTCTGT
 TTGTTTTAATAAAGTAATATCCGTAAAATACGTCAAGTGAAGAAAAATCA
 TCGCTGAAAAGTAAAAGGTATPATTTAAATTTATAGTAAATGAGTAATTA
 30 GACCTCCGACTAGTGGGCTATCATTGAACCGAAGCCTTGAACACTGTTGA
 ACACGCCCATGTTTCTTCTGTTTCGTTAGGATTAATATGCCAGCCATAA
 AGGTATTCCAAGCCGGTAAGAGGATACCGTACATTAGCCCAATAAAAAGTC
 CTATGGCCCAACTATATATATATTTGTAATTGTAGATAGCCCGAATATAA
 GAATTGTATATAGTATAAAGCCACTAAAAATAACTCCATACATAAACCTT
 35 TGCTATTATTGTGCGATGATTTTGATAAAAATAACATAGAGAAAGCACAGC
 CTATGCCACCAATAATGATTGCTACTGTATATTCAACGGTTGATACTTTCA
 CAACTTGCGTTGCATATTTTGGAAGAATAGGTACAAGTGCTGCTATAGCTG
 CTCTTTGTAACAAGATACCCGAAATAGAATAAGATGACG

40

Sequence 3407

step.1002f08.cons.ok

TAACGTGTGTCTCGCTATCAGTAATTACTGCAAAAGGTGTCACCTCAGAAC
 CAGTTCCTGATGTCGTTGGTATACATATAAATTTGCGTTTTTAGGTTTG
 45 TAATTTTATAAGTACGTTTACGAATATCTAAGAACTTTTGTTTTGCCCCAA
 AAAATGAAGTTTCTGGATGCTCAAAGAACATCCATATTGCTTTGGCTGCAT
 CCATTGCCGAACCGCCACCGAGTGCAATAATAGTATTAGGTTGGAAATTTA
 TAAACATTTCTAACCCCTTATAGACTGTATGAGTTGATGGATTAGGTTCAA
 CTTCTGTTAAACACTTTGATTTGTGGTTGGTTTCTCGGCGTCTCAGCACTT
 50 GTTCAACTATATCAGTATAACCAATATTAACCATTCCTGGATCACAACTA
 TCATTACAGCTTCAACATTATCCATCTCTGTCAAATACATAACTGAATTTT
 CTTCAAAATAAACTTTAGGTGGGAGTTTAAACCATTGTCATATTATTACGAC
 GTTTTGCTATTGTTTTAATATTTAATAAGTCTACTGCGCTTACATTATGAG
 AAATAGAATTTCTACCATATGAACCACAACCTAACGTGAGTGAAGGAATGA
 55 GTTCATTATACATATTTCCAATTCCTCCGACAGCAGAAGGTGTATTACCA
 ATACACGGCAAGCTTTCAATTTTAGACCGAATTTTGTGTGAATTGACTAT
 CCTCGGTGTGAATTACAGCAGTGTGACCTAAACCACCAATTTTAATATGT
 CTTACAAAATTTGTAGTGCATGTCTGTTGATTTGCAGTTACCATGCGA
 GTACAGGTGATAATTTTTCACGTGATAAAGGATAATCTTTTCCAATTCAT

CAATTTCTGCGACTAATAATTTTGTGTTTTCTGGAACACTAATTCCTGACA
ATTTTCGCTATATCTACAGCAGATTTACCAACTATATCAGGTTTAACTGCAG
TTTTATCTTCATTCATGATGGCATCTTCTAATTGTTGTAGTTCATTTTAT
TAACAAAATATGTTTGGTGTAATTTGAATTTCTTTAACGACGTCAGTGTATA
5 CTTCTTTATCAACAACCATGACTTGTTCAGAAGCACAAATCATACCATTAT
CAAAAGTTTTAGAACCAATGATATCATTAAACAGCACGTTTGATATGAGCAG
TTTTTTTCAATATAAGTAGGAACATTACCTGGACCGACTCCTAATGCAGGTT
TACCTGTGCAATATGCGGACTTTACCATTCCAGAGCCTCCAGTCGCTAGAA
CTAAAGCAATATCTTTATGATTCATTAATTGTTTAGTTGCCTCAATTGATG
10 GCACTTCTATCCATTGAATACAATCTTTAGGAGCACCTGCTTTTGTGTCAG
CTTCTAAAAATGACTTTAGCAGCATATTTTGATGATTTGTTGTGCACTTGGAT
GAAATGCGAAAAATAATTGGATTACCTGTTTTAATAGCAATCATTTGCTTTGA
AAATAGTTGTAGATGTAGGGTTGGTCGTTGGTGTCACTCCACAAATCACGC
CTATAGGTTTCAGCTACATACGTTAATCCTTTTTGTTTATCTTCACCTATGA
15 TACCTACTGTTTTATTGTCCTTTGATTGAATTCATATGTACTCTGAGGCAT
ATAAATTTTTGTAGACTTTGTCCTCATAAATACCTCTACCTGTTTCGTCGT
AAGCTAGTTTAGCTAAATGCATATGCTGATCAACAGCAGCCATGCTCATCT
GATGTACAATGTCTAATCTCATGTTGTGATTTTTTAGATAGTTCTTTTA
GAGCTTCTTGTCCTTTTTCTGCTAATGAATCAATCATTTGTGTGACTTCAT
20 CTTTGTGTTGATTCATATGTATTTTTTTTAGTTACAGATAACATATACACC
ACTCCTTAGATATTGTGAAATAATTCACAAACATTTAGTACAGCTCTTTC
CAAATTTAAAGAAAATACAGACTATTAAAGAGTCAATTGTGATGTTTATAT
GAACACATTTTTAACCTATATATTTAGTCGGGGTTGAGGTGTAAGACTTGA
TGCAATGGGAACAAGACGTATTTGGTAGTTAAAAAATCTATTTTTTAAAG
25 AAGTAAATTTTTAAAAAATTTATTGTAAGTATTCTTACTAAAAATAATGAGG
GGTGTGTAATGAAATTATCGTCTCACCCGACTAACGATGTTATCAGTTTA
TAAATAAAAAACTGAGACAATCATTTATGTCTCAGTAGCTCATGCTTAA
AAATAAATTTGATAAAGAATAAGTGAAATTTATATTAATTATTTTGTTCAGT
TGCAGTAAGAAGTTGTATCGCTTTGGGTAATACTTGAATCTCAATTGGTGT
30 TTCTAAATTAATTTGCCATCAATATCCACTTTTCACTAGGGTTTGTGTA
GAGTGTGATGTGCTTACCTGATATGTGATCAATACCTTGTGTGATTTTCGTT
CCAATTCATACTATCACGTTTTTTTAAATATATCATTTCAATATATTTAGTGT
TTGATCATTAATAACAAATGTGTTTGCTCTTCCATCTTGTGGCGATAAATC
GGTTAGCGGAATTTGTCCACCACCTATATTGGGACCGTTTGCTATTAAACAT
35 CATCGAAGTATTGCCTTCTTTTGTTCACCATCAACAGTCAATGAGAAATC
AAATTTAACAGGATTTAATAACGTTTAAACGTTTGATCCAATATAGCTGAA
TTTACCGAATATATCTTTAGAACCATCTTGAACATTCTCTGCATTTTGTAC
TATTAAGCCAAGTCCAACGAAATTAAGTACATATAAGTCGTTTCACTTTTAA
CACATCATATGATTACGATGTGATGTTAATAATTGCTCACTAGCTGTTTTT
40 AAAATTAGGGTGCAGTTGAAGTGTTTTTGTAAATCGTTAAAGGTACCACC
TGGTATTACACCGATTGGTAAATTTAACTGATACTGCATAACGCCATTTAC
TAGTTCATTAAGTGTACCATCTCCACCTAAAAATAAAACGCTTACATC
AGAGCTATAATTTTCATTTTTAATAGATTTACAATATTTAATAATATCGCC
TTTTATTTTCACTGAGTTGAAGAGAGAGGTGTTTACACATTGAACTAAGAGA
45 TTTTGCAACTTCTCCTATGCCATTATGTATATCTTTCAAACCGCTATGTTT
ATTCAAACTTGGCACGGTACACCATTAAAACGTTTAGCGAATGATATGAAA
GTGGTACGTATGCCGGGTACTACGACACCAAAGTATAAGCGTAATTTTAAT
CGTGAAACATCACGTTGGGATTATTTAATTTCCGCCAAATAGATATTTCAACT
GAAATATTTAGAAGTGCTTTTTGGATGGATGAAGAAAGAAATTTAGAGATA
50 GGTATATCCAAGAAATGATGTATTAGTTAATAGAGCCAATGATCAAGAGTAT
TTAGATGAAATTTAGAATCACTTAAATTTACCTAGTGATAAAAAAGGTATT
ATGTATGCTCCGACATGGAGAGACGATGAATTTGTGAGTAAAGGAAAAATAT
TTGTTTGAATTTAAAAATTGATTTAGACAACCTTTATAAAGAACTCGGAGAT
GATTATGTGATTTTATTACGCATGCATTATCTCATTTCTAACGCACCTTGAT
55 TTATCTGGTTATGAAAATTTTGAATTTGATGTTTCAAACCTATAATGACGTC
TCTGAATTATTTTAAATAAG

Sequence 3408

step.1002f10.cons.ok
CGCCACTGTGCTCTAAAGATATACGTGTCGCGCAACTTGTTCCTGTCGTGC
TTATTATGATTAGTGTGATATTCGTAATTTATAGAAGAGTTAAATACCAAC
CAATAAAATATGAAAATTCAGGACCATTAAACATGGCCGATTAAAAAGGCTA
5 AGTGATTATTTTGAGAAGATTAAAGTAAAGAGAAGCATCATCAAGTGAATCC
TTTATGGCATATATATAGAAAAATTAAATTTCGTAAAAGTGCTAAAGCAAAC
GTTGATTATAGAAAATTGCAAGATTTGTTCCCTAGTATGAAACTTAAAAATAA
AATATATAAAAAAGCTTTTAAAAATGGATGTCGGGAATCATACTTCATTGTC
10 ATATAAAGTGTTGCCTGATTTGTTTATCCAGAATACATTTCCGTTGGCAA
GAATACAGTCATTGGTTATAATACAACAATATTAACACGAAGTACTTGT
TGATGAGTGGAGAGTAGGAAAAGTCATTATAGGCGATTACACTTTAATAGG
TGCAAAATACGACAATATTACCAGGAATAACCATAGGAAATCATGTTAAAAAT
TGGTGC GGGTACGGTTGTGCTAAAGATGTTCCCGATTACAGTTTTCATT
TGGTAATCCTATGCAAATACAATTAGATTCAGGAGGTGACAATGAATGGCA
15 CAAAAAGAAAATACATCATTCCAATGATTTTGTATGAGGCTTTTATCGT
AAAATGGCTACCCAAAAGTGGCGACAACAGGACTACAAGAAGGCGGCAGAA
TATTATGAAAAAGTGCTTGAGTTGTACCTAAGGATTTTGATATTCAACAA
CACTATGCTCAATGTTTAGTTAAATTAATATCGGCAAAAAAGCAGAACGT
CTATTTTATGAGAAATATTGTTAAAGACTTTCATGTAGCAGAAAGTTCTAT
20 GAACTAAGTCAATTAAATATAGAATTAAACGAGCCAAAACAGGCATTCTTG
TTTGGTATTAATTATGTAATATTGTCAGAGGATAAAGATTTTAGAGAAATG
CTTGAAAAAACATTGGAAGTCACGTATACAAATGAACAAAAAATTGAATTA
GAAGCACAAATTGTTTCAACACAACCTTTATTTCAATTTCTCTTTTCGCAA
GGTAGGTTAGAAGAAGCCCGAACATATATTTGAATCAATCTTACGAGATA
25 CAACAGCATAGGGTGATTAGGAATTTACTTGCAATGTGTATTTTGTATCTA
GGTGAGTATGATAGCGCCAAAGCAATGTTTGAAGAACCTTTAAAGGAAGAT
AATTCAGACGTGCATGCACCTTTGTCACTACACATTATTACTTTATAATAAA
AAAGAAACAGAAAAATATCAAAAATATCTTAAATACTTAATAAAGTAGTA
CCACTAAATGACGACGAAACCTTTAAATTAGGAATCGTATTGAGTTATTTA
30 AAACAGTATCGTGCTTCTCAAAATTTACTTTATCCACTTTATAAAAAAGGT
AAATTTGTCTCTATTCAAATGTATAATGCATTGAGTTTCAATTTTATTAC
CTAGGAAATAAAGACGAAAGTATTGAGATGTGGAACAAGCTCACTCAAAT
TCTGAAGTTGATGTTGGTTATGCACCTTGGGTAATTGAGGAAAGTAAACG
GTATTTGAATCACGAGTGTTACCATTATTACTAGATGATAATAATCATTAT
35 CGACTTTACGGTATTTTTTTTACTTCATCAATTAAATGGAAAAGAAATACTA
ATGACTGAAGATATTTGGTCAATTTCTTGAATCAATGAATGACTATGAGAAA
CTTTATCTCACATATTTGGTACAAGGACTCACTCAATAAATTAGATTTT
ATACACAGAGGATTGCAAGGTTGTATAATTTTAAAGAAATTCAAATATAAC
ACGTCTTTATTTACAGATTGGATTAATCAAGCAGAAATGATTATAGCTGAA
40 AATGTAGATTTAGTAGATGTCGATAGATATGTAGCTGCATTTGTTTACCTA
TCGTATCGTCGTCTAGCCAACCACTTACCAAGAGGCAATTGATGGACGAT
TTTAATGTTTCTAGATACAACTGAATAAAGCAATTGAATTTATATTGAGC
ATATAAATTTATGAAATGAAAAGTATGATATATAATGCGCATAATGATTAA
TAAGTAGGAGGCATTTAAATGACTGAAGTAGATTTTGATGTAGCAATAAT
45 CGGTGCAGGTCTGCCGGTATGACAGCAGCAGTATATGCATCTCGTGCCAA
TTTAAAAACTGTCATGATTGAACCGGTATGCCAGGCGGTCAAATGGCAAA
CACTGAAGAAGTAGAGAATTTTCCAGGATTTGAGATGATCACAGGTCCTGA
CTTATCTACTAAATGTTTGAACATGCTAAAAAATTGGTGC GGGAATACCA
ATATGGCGATATTAATCTGTTGAAGATAAAGGCGACTATAAAGTTATCAA
50 TTTAGGGAATATAGTTGTTGATGTAAATGTGTGTACAGTATTGTTTAAAC
GTATTTTCATAACGTAATAACAATGTGTCAATCATCGGTGTTGCAACGTTT
AATGCTCTTCCAATCGCTTGAATTATCAAAAGTACGATAATAATCTTCACTC
GGCATGCGTGGTATATGTATGACTCCTTGTTCATCAGTATCAACATGTTTG
TATGGTACGGCAGAAAAATCAAAATATGCACCTTGATCGTCCGGATTAGAA
55 AACGGATCGATTTAAATTTGCTGTATATCGCACATAAAGTAGATACTCTTGA
TGAATAGCTGTTAAATTTTTTAAAGTTTTCATATCTACTTCGCGCATGGTC
TCATAACGTATAGGGTAGTTTCTTTCACCATAAACTTTAGAAGATTGACC
GAAGGTACCTTTAATTTTTTTAATATCATCATCATTTCTTGCCACATTAAT
CGCATTTCTGTATTAAGGTCATTGTGATTGGACCTCTGGAATAGCTTA

TATACATATACTGGTACTTTCGTCCCTTCAAATACCGCCTTTAATGAAAAT
 TGATTCTATAAACAATGGTGGGTGTACATAAAGTGAACCTATTATGTATCTCC
 GCATGTAGTGGTGTGTCCATCGTTGTTAATTCATATTCAAATAGTCAAAT
 AAAGACTTAAGCTTACACAACGTCATAGACTGAGATTGAGTCGAACCTACG
 5 AATAATTTTGATTTAACTCGTGTGGTTAGGACACAATGTGGTTGTGCTTTA
 TCAAATATTCGGGTATCGCCTAGATAAGTGGAAAATGAAATCACTTCACCT
 TCACATTGAACATCTGATAGTAATTGCTTAACAAGCATATGTGATCCTAAT
 GTTGGTGAGACCAAGATGATTGCTTAATACGCTTTAATGTGGACTTAGAT
 AATTGCTGTAATATCGGTCGATACGCATCGGCAGTACATGCTAAAATCACC
 10 ACGTCATAATATTTCAGTAATATCTTTAACATCTTTAAAAAATGTCTAACC
 GTAAACTTACCTGAAAAACACTGATGTGCATCATTTTGAGTCATTACTGAA
 AAAAAAGCGTCACGTTGATAAGCATCAAAGACTCTCTTAGATTGGTTGAT
 GCATGAACGCGACTCACCATATCAACTATATGTTCTCCATGTAAATGACAT
 AGTCGAGCAAGTTGAATAGCGACCGGTCCACTACCAACCATTAAATATTTTA
 15 GACATGTTCAACACCAACTTTTGATGCCACTTTACAATAAAGTGCAATATC
 AAAAATCTGCTTCGGTCGTGTTTGATGTTCTTCACATCGCCACTGTGTTTC
 GTCAGTTGCTTGAGATGGATAATTAAACAATGCTTTTATATCATCACCCTA
 TCGCATAGCTACAACAACCTTCATCATTCGTCAATGTGTAAGTTTCATTCAA
 AATGTCATATTTGATAGGTATCGTCGAAGTGAATATAATATGTGTACAGA
 20 TTGG

Sequence 3409

step.1002f12.cons.ok

25 AAGATGTAAAAAACAATAATGATTATTTAGATAAAAATTTTACCTGAAGCTT
 ATGCACCTGTACGTGAGGGGTCAAAGAGAGTATTTAATATGATTCCTTATA
 AAGTACAAGTAATGGGTGGTATTGCTATACATAAAGGTGATATTGCAGAAA
 TGAGAACAGGTGAAGGGAAAACATTGACTGCAACCATGCCGACGTATTTGA
 ATGCTTTAGCTGGTAGAGGTGTACATGTTATTACAGTCAATGAATATCTAT
 30 CAAGTTCACAAAGTGAAGAAATGGCTGAACATATATAACTATCTTGGCTTAA
 CTGTAGGTTTGAACCTTAAATAGTAAGTCAACTGAAGAAAACGTGAGGCTT
 ACGCACAAAGATATCACCTTATAGTACGAATAATGAACCTGGGTTTGATTATC
 TTAGAGATAATATGGTGAACATATGCTGAAGAGAGAGTAATGCGTCCTCTAC
 ATTTTGCAATTATTGATGAGGTCGATTCCATATTGATCGACGAAGCAAGAA
 35 CACCTTTAATTATTTCTGGTGAAGCGGAAAAATCTACTTCTTTATATACTC
 AAGCAAAATGTTTTTGCAAAAATGCTTAAAGCGGAAGATGATTATAATATG
 ATGAAAAAACCAAGCTGTACATCTTACAGAACAAAGGTGCAGATAAAGCTG
 AACGTATGTTCAAAGTAGATAATCTTTATGATGTTCAAAATGTGGAAGTGA
 TTAGTCAATATATACAGCTTTAAGAGCTCATGTTACTTTGCAACGCGATG
 40 TTGATTACATGGTCGTTGACGGTGAAGTATTAATTGTTGACCAATTTACTG
 GACGTACAATGCCTGGACGTCGTTTCTGTAAGGTTTACACCAAGCAATTG
 AGGCTAAAGAAGGTGTAGCAATTCAAAATGAGTCTAAAACGATGGCATCCA
 TTACTTTCCAAAACATTTTCAGAATGTATAAAGTTAGCGGGGATGACTG
 GTACACGCAAAACCGAAGAGGAAGAATTTCTGTAATATCTATAATATGACAG
 45 TTACCCAAATTCCAACAAACAAACCTGTTCAACGTAAAGATAAATTCAGACT
 TAATTTATATTAGTCAAAAAGGAAAGTTTGATGCGGTAGTTGAAGATGTTG
 TAGAAAAACATAAAAAAGGACAACCCGCTTACTAGGTACTGTTGCTGTTG
 AGACTTCTGAATATATTTCAAATTTACTAAAAAACGTGGTGTGACAGATG
 ACGTATTAAACGCTAAAAATCATGAACGCGAAGCTGAAATCGTTTCAAACG
 50 CGGGGCAAAAAGGTGCAGTTACAATTGCCACAAATATGGCTGGACGTGGAA
 CAGATATTAAACTTGGTGATGGTGTGAAGAGTTAGGTGGACTTGGTGTTA
 TTGGTACTGAGCGTCATGAATCAAGACGTATTGATGATCAATTACGTGGAC
 GTTCAGGACGCCAAGGTGATAGAGGAGATAGTCGTTTACCTATCTTTAC
 AAGATGAATTAATGGTACGTTTGGTTTCAAGACGCTTACAGAAAATGATGA
 55 ACCGTTTAGGAATGGATGATTCAACGCCAATCGAGTCGAAAATGGTATCTC
 GAGCTGTAGAATCAGCTCAAAAACGAGTAGAAGGTAAATACTTTGACGCGC
 GTAAACGTATTCTAGAATACGATGAAGTTTACGTAAGCAACGTGAAATTA
 TTTATAATGAGCGTAATGAAATCATTGATAGTGAAGAAAGTTCTCAAGTCG
 TTAACGCGATGTTACGTTCTACATTGCAACGTGCGATTAATCATTTTATTA

ATGAAGAAGACGATAATCCTGACTACACGCCATTTATCAATTACGTTAATG
 ATGTGTTCTTGCTGAATTATTAATGTTAAATGCTATCCATGTGCCATTACA
 CACCGGTAAAAATTGAAGAAATGACACGTGTACTTCGTGAAAAATAAGCATA
 ATATATTTAAGCGTCTGGATTATCATCCCCAGACGCTTTTAAATGTTTAG
 5 AGAATTATCAAAATTTTATTTGAATGCTTTTCGCTAAAACAAAGATAAACATA
 TCTTTTCGAGGAGTTTATCTAGTATACGTACATTCTTCCCTCAACAAGCTA
 AAAACATAACTATCTGAGAATTCCTTGTAAATCGTTCATTACTTCTCAAT
 ATACCTTTGCGTGAACCCAAGTCTAATTGGTATAGCTCTGCTTTTCTTATT
 TTTAGTAGACATTTCGTATTTCTATACGATTAATATCATACACTTCATATAC
 10 ATAGCGAATTAAACGCTTTAGTACATTTAGTCATAATACCTTTCTTTTGAAA
 GTCTTCAGCTAAATAATAACCAATTGAAGTTGTTTATTAACATAATCTAA
 GTAATGCAATCCTATGACTCCAATCAATTCCTTTATTACTCCATATTCCACA
 ATGAAATCCATTACCATCGATAAAATGTTGCAACGCCGAATGGATAAAGTG
 TTTACTATCTTCAACTTTCTTCGTATGTTCAACAAAAGGCAGAAATTCAGC
 15 TAAATAGTCACGATTGCTATCTACTAATTTAAATAACTGTTTCGGCTTCTCG
 TTCTTCTAATATTTTAAATTTAATATTTTATTTATTTGATAACTAAACAT
 ACAAATCATCCTTTCTTAGAATTTCTGTCAATTTAGTATACGCGAGTAT
 AGGATTATGTCATAAACTTGATTACTAGAATTCGATATTCTTTAAGTTAC
 AAGTGTGTATGTAGTATTTCTCTTGCTTCTATTTGCTTATACTTACAACCA
 20 ATATTAGAGTGAAGGCTTAACTTAAGATAGATAATCATCAACAACACAA
 ACGTTAAAAACACTGAGAAAACCTTTTTTAAATATACTTCTCAATATTGAT
 TAGTAAAAATCCAAAATATGAATAATAAGTAGACATTAAACAATCTTTATTT
 TATTCTATAAAATAATAAAATAAAATATTAGATTAAATACATTATTTTATTT
 ATTTCTAATTTAAGTTGAAATCGTATGTAATCGGAGGAAACCATAATGATA
 25 TATTCTATCGGTAAGAATTTAGGTAATAAATTAACAGGTATAGAACAAGCT
 ATGATCAATAGATTAAAGCTATTTAAAGATAATTTAGTCCCAAATAAACTC
 ATATTACATCTTGGTCACCACGTTTATATATGCATGCACATTTCGTTAAAC
 ATCGATTCAAAAGATATTTTCAGTCTTTACGATTTTCTACAAGATAGTATT
 AACTTTGAGAAAAACATATTGATTGGATAAATTATTGGCAAAATATATGT
 30 AATTATACCTTAAATTCGTTGAAAATACGAATGATATTAAATATACGAT
 AACGACACATATAAAATGTATGTGCATTTTGTGATTCAAATTATCAAACCT
 TTAGACTATATTAACCATTTTGTATATAACAACGTAAATTCGAAGAGAT
 TTTTACGATACAAGAGGCTTTTAAAGTTGTAGTAGAATTTTAACCTCTCAA
 C

Sequence 3410
 step.1002g01.cons.ok
 TAACCGTAGGTATAGTGGTTAGCTTTACATAATTCAATGACTTTTGTGTT
 40 ACGGTATCATTTTTATAGGTTTGTTTTTCCATCGGTAATATGTTGATTTA
 GGTATGTTTAAACTTCTAGTATCAATTTGATTGAATATTTTACTTTTAAAT
 TGAATCCACTAAATCTATGACTACTGTTGGTACCACCTTCTTTCCAATGCCT
 TGTACTTTTTTAAATATCCAATTCATATCTTTTCTCTTATTTTCTAATT
 TTAATTGTTCTACTTCTGACAGCTCTTCTAATCCTTTTACCGTAGGTATATT
 45 GTTTACCAACTTGTGTGAAAATCTATAACTTTCCCCATTTTCGATACCATC
 GCCACCAAGTTTTCACTTGCCTCTATTTCTAATATTTAATTCCTTTCATAA
 TTTCTTTTGTGAAAATCCTGCTGCTTTTCATTTCAACTGCTTTTACTTTG
 TTTCTACTGAATAAGAAGCTCTTTTCATAGAAAAAACCTCCGTATGATTTC
 ATTTTAATATGAATTCAACGAAAGTGTTTTTATATAATTCCCACAAATTGG
 50 GGTCACTTAATGAATTTCTGATTGTAGGCTTTTATTTTATATATT
 CATTAATTTCTCTTCTTGACGGATACGTACAATTAATAAATAAGCATATGG
 AACTAATAAGAGTGTGTGTATGTAGCATTGTGTAGTAGTAATACACCAAT
 TAATTCGGGAATGATATTTAAAAAATAGTTTGGATGTTTCGTTACTTTATA
 TAACCTGACTTAATAATAGGGTGAATGGGCAGTATAAATAATTCAATGT
 55 CCAAATACGACCTAATGCTTAAATAACTATAAATAGCATGATATAAGCGAT
 GATCAATATGATTAAAGCCAATGCCGTTAAGTAGACTAAATGTATCCTTACG
 AATGAATGCTTCTATAGCTGCACTCATGTAATTAATACATGCGTAATGGC
 TAGATATTTTGAATTTTTCACACCATATTCCACCGCGCCCTCTACCTTTAG
 CTGTTTTGCGTGTGTCATAGATATCTTTAAGCTGATGAGTCGAATACAGAA

AAAGATAAAATAAAATAGTTAAAAATCATGATGTCCTCCATTGTTTGTATAT
 ATATATTCCTTTGCTCGTACAATACAAAGTATTTTATATAGAAAAATGTTTC
 AAATGCCATTAACTATTTATTAACCTTAGCACCAACTCATAACTCTCTAA
 CACCTCACATATTAAGTTTCAGTTTTCGGATAAATTAATACTTTTAAGGAT
 5 ATTAAGCGCTTACATTGATGTGATATATTTTTTTTAACGAAGATGATAGGG
 GTGTTAGGGATGAACTTTTTTGATATTCATAAAATGCCAAACAAAGGGATA
 CCATTAGCTGTACAACGCAAATTATGGCTCAGAACTTTATGCAAGCGTTT
 TTTGTCTGATCTCTTTGTTTACATGGCGATGTATTTAATTCGAAACAATTTT
 AAAGCGGCACAACCGTTATTAAGAAGAAATCGGATTAACAACATTAGAA
 10 CTAGGTTATATAGGATTAGCGTTTAGTATTACTTACGGTTTAGGAAAAACA
 TTAATCATTTGGGATTGTCGTTGATGGCGTAATACGAAACGTATTATTTCCTTC
 TTATTAATATTATCTGCGATTACAGTACTTATTATGGGATTGTATTAAAGT
 TATTTTCGGTTCTGTGATGGGGCTATTAATTGTATTGTGGGGGCTTAACGGT
 ATATTTCAATCTGTGGGTGGGCTGCAAGTTACTCAACGATTTCAAGGTGG
 15 GCGCCTCGAACAAGCGCGGTCGTTATTTAGGCTTTTGGGAATACATCACAT
 AACATTGGTGGTGCTATTGCTGGTGGTGTGCGACTTTGGGGCGCGAATACA
 TTTTTCACGGTAATGTGGTTGGAATGTTTATTTTCCTTCCTCATCGCT
 TTAATCATTTGGGATTGTGACATTATTTATTGGTAAAGATGATCCAGAGGAA
 TTAGGTTGGAATCGTGCCGAAGAAATTTGGGAAGAGCCTATCGACCAAGAA
 20 AACATTGATTCTCAAGGTATGACTAAATGGGATATCTTTAAAAATATATC
 CTTGGAAATCCTGTGATTTGGATTGTGTATCTCTAATGTTTTGTATAT
 ATCGTGCGTATTGGTATTGATAACTGGGCACCGCTATACGTATCAGAGCAT
 TTACATTTTAATAAAGGTGATGCGGTGAATACTATTTTTTACTTTGAAATA
 GGTGCATTAGTAGCTAGTTTATTGTGGGGCTATATCTCAGATTTATTAATA
 25 GGTGCGTTCGCGATTGTAGCGATTGGATGTATGTTTATGATCACCTTTGTT
 GTACTCTTTTATACCAATGCAACAAGCGTGACAATGGTCAATATTTCTCTA
 TTTGCATTAGGCGCTTTAATCTTCGGTCCACAGTTACTCATTTGGTGTATCT
 CTGACTGGCTTTGTTTCCTAAAAATGCAATTAGTGTGCTAACGGTATGACA
 GGTTCAATTTGCATATCTATTCGGGGATTCAATGGCTAAAGTGGGTCTGGCT
 30 GCAATCGCTGATCCAACACGTAATGGTTTAAATATTTTTGGGTATACGTTG
 AGTGGTTGGACAGATGTCCTTATTGTATTCTATGTAGCTTTATCTTAGGA
 ATGATATTATTAGCCATTGTTGCTTATTACGAAGAAAAGAAAATTAGAAAA
 TTAATAAATTAAGATTAAGTGAATTTAAATATATACTCCCTACTATAAATT
 TTATCATTTATGGTAGGGAGCTTTTTCGTGATTGATTGCACATTTTGATT
 35 GCTTATTTAATCACTTGGTGACATATGTAAATATTTTTTAAATGATAGC
 AAAACATTTTATACTCAGAAAAACCTACTTTTTCAGCAATTTTCAT

Sequence 3411

40 step.1002g02.cons.ok
 TTAAGTTGTGTAAGTTATCTACTTGTTCAATCCATTTCGCTATCTATCGCT
 TTCAAAATTGATTTTTGAATGAAACGAAGTTTAAATATGAATCAGCAGCA
 ACTTCTAAACGATTGTTAAATGTTGAGTAAATGTTGTATTAAAGAAATTGT
 ATGATTTCTTCATCATTTTGCATATTAATATTTGATACATCTTCATCGAAG
 45 ACAAACCTTAAGTTTTTCGTATATATAATTCACAAGTGCACGTTCACTACTT
 AAGTCAAGATTTTAAACGTCTTTTGTAACACATCTCGTGCAAGCTGTTCA
 AAATTAATAATCATCAAAATCGCTTGCTTCAAGTATGTGATTACGTTTCAGCA
 TAAATTTTATCTCGTTGAACACTAATACTTTTTCGAATTCATTGCCATT
 TCTCTATTTTTCATAGCAGTCTCTTCAGATACACGTTGCGCTTTATTAACA
 50 ATTGACTTTACACGTTTTTTAAAGAGTGCCTACTTTCTAGTTTAGATGCA
 TCCATCGTTTGGAGGTTTTTATTTTCTGCCAAGTTAGAGTTACTCCAACGT
 TTTACTAAATCATCAAGTGATACAAAACTCTGTGAATATCCAGGATCT
 CCTTGGCGACCTGAGCGTCCCTCTTAATTGACGATCAACACGGCTATTATCC
 ATATGTTTCATTAATAATCACTGCTAAGCCACCGATATCATGAACCTCTTTT
 55 GATAACTTTATATCAGTTCCACGCCCTGCCATACTTGTAGCAACAGTAACT
 GCAGATAATTGTCCCGCTCAGCAATCATTTGTGCCTCTTTAGCTACATTT
 TGAGCGATTAATAAATTTGTTGGGTATATCAGCTTTAAATAACTCAGCTGAA
 AAATATTTCTGCCGCTTCGTCAGTACGTGTAATTAGTAACACAGGTTGTTGA
 GTTTCATGTATACCAATCACTGTCTTTAAATTTGCATCGTTCTTTTGTCA

CCATTAGCAAATACTCTATCAGGTCTATCATCTCGTTCAATCGGACTGTGA
GT

5 Sequence 3412
 step.1002g03.cons.ok
 GTATTCGGTAATATCATCTGTAAGATTTTAAATTTCTTGTCTTGCATTTT
 CATTGCAACAGTGTGATGGCCAAAATTAAACATAGCTCTATCATTATCACT
 ATCACCAATAACTAAAGTTTCTTCTTGCTTAATACCATAGTGATCAATCAT
10 TTCCTTGATACCTGTACCCTTATCTGTATGATATGGCATCGTTTCAGCAT
 ATAACGCGAAGAATTAGATACGGTAATACCTAAATTTGATTGGTTTCAAC
 AAGCTGGTTGCGAAAATCAGTTATTTTATCTAAGTTGGGACTAAATAAATA
 TATTTTAGAAAAGTGTGCATCAGGTAAGGTATCTCGCCAATCTATTTTCC
 TTTAAATGGCTCTCTTCGCGATGACCACTCACTTTGACTTACAGCGTCAGG
15 TGGCTCTATAGTGGAATCATTTCTTTTCATCCAATCTTCATCTTCTTTAAG
 AGATATACGATTACTTTCAAAGGAAAAAATTCATAATAAATATGTTGTTT
 TTTAGCCAATTCCACAATTTTCTGTACTCGTTCTAAAGTTAAACTATGTCT
 AAACAAATTATCTCCATGAAATTCACCTGAAGTTCCATTGGAACGTATAAT
 ACCATCTACAGTGAATCCATCAGGAACAAGCTGACTGATTTCTGAATAAGA
20 TCTCCAGTAGCAAGGAAAACTTATAATTTTGTCTCTCAATTCATTAAT
 CACTTGTTTAGTATATTCCGAGGCTTTATTATTTTCATGTAAAATCGTTCC
 ATCCATATCTAAAAAATTGCTTTTATATTTTCATATTATCCTCCATTCA
 AGACTCTAACTTATACACTTAAATTTGTGTCACTTATCATACTAAAGTGCAA
 CGGCTTTGATTACACTTTGAACATACTATAACGTTGTCTTCTCTTATTAC
25 AATGACATCCCTATCAAAAAAATTTGTATCCGATAAAGATCTCATAACATC
 TTAAATATTAACCTTGTAAATATTTAAAAGTCTTTAAGTAAATCTTTTTC
 AAAAATAACGGCATTTCTTCTATTATATGTGTTACAAGTATGACAGAAAGT
 TGTGAATCATTTAAATATGCTTACCAATGTTTAAATAGTTGTCTCTTT
 ATATAAAATCCATTCCAGTTGTAGGTTTATCCAATACTAGTTAATTTAGGG
30 TACTCATTAAAGTCATTTAACAACGGCATTAAACCGCTTCGTTAAATTA
 TTGAAGGATAAAATCTAAAAATTAATTCAAAAATAAAGAGTAGCCATTTAGGC
 TACTCTAAATTATAAACATTTTACTCTACTGTAACCGATTGGGCTAAGTTA
 CGAGGTTTGTCAACATCTAAATCTCGTTGTAACGCAGCATAATATGAGATT
 AATTGCATTGTCACTACAGATACTAAAGGAGTTAATAATTCATGTACTTGT
35 GGAATCACGTATGTGTCCTCTTCTTTATTCAAACCTTCCATTGAAATCATA
 CAAGGATATGCACCAGTGTCTACTACTTCTTTTCATATTTCCACGAATTGAT
 AGATTAACGTTTCTTGTGTAGCTAAACCTATAACAGGTGTGCCATCTTCA
 ATCAAAGCGAATTGTTCCGTGCTTTAATTCCCCACCTGCAAATCCTTCAGCT
 TGAATATAAGAAATTTCTTTTAATTTTAATGCACCTTCTAACTAACATTA
40 TAATCAATGTTCGTCCAATGAAGAATGCATTACGAGTAGTTTTTAAGAAA
 TCCGTTGCAATTTGCTCCATCTTAGGTGCATCGTCAACAATTGTTTCAATA
 GCTGTAGTAACCTTAGCTAGTTCTCTTAATAAATCAACATCGGTTTCACGA
 CCATGATTTTGTAGCAACAATTTGAGATAAGATAGATAAAACAGCAATTTGC
 GCTGTATATGCTTTTGTAGATGCGACTGCAATCTCAGGTCCAGCATGTAAA
45 AGTAATGTATGATCCGCTTCACGTGATAATGTTGAACCAGCAACATTAGTA
 ATTGTTAATGATTTGTGACCTAACTTATTTGTTTCAACTAATACAGCACGA
 CTATCAGCTGTTTCACCAGATTGTGAAATATAAATAAATAGTGGTTTTC
 GAAAGAAGTGGCATATTATATACAAATTCAGAAGCTACATGAACCTCAGTA
 GGTACACCTGCCCATTTTCAATAAAATCTTTTACCAACCAATCCAGCATGA
50 TAAGTAGTACCAGCTGCAACGATGTAAATACGATCAGCATCTGCTACATCA
 TTAATAATCTCTGATCGATTTTTAAATTACCTTTTTCATCTTGATATTCT
 TGAATAATGCGACGCATCACTGCAGGCTGTTCAATGAATTTCTTTTAAATG
 TAATGATCATATACGCCTTTTCTGCATCAGCAGCATCTATTTCTGCCGTA
 TACGTATCAGGTTGTTGAATGTGCCCCCTCAAGATCTTAAATTTCTACTGTG
55 TCTCGCTTAACAATAACTATTTTCATGGTCATGTATCTCTTGTATTGGTTT
 GTAGTTTGTAAATTTGCTAGAGCATCAGAAGCAATAACATTGAAACCTTCA
 CCTACACCTACTAAAAGCGGAGACTTGTTTTTAGCCACATAAATAGTATCT
 TTATCATTATCATCTAATAATCCTAAAGCATATGAACCATGTAATAATTTA
 ACTACTTTTGTAAATGCATCTTCTGTAGCTAATCCTTGTCTAGAAAAATAA

5 TCTACTAATTGTACAATAACTTCCGTATCAGTTTCTGATGAAAAAGTGACA
 TCAGATAAATATTACAGCTTTTAATTCTTCATAATTTTCAATTACACCATTA
 TGAACCTAATGTAAACGTCACAGATGTTGACTGGTGTGGATGTGAATTTTCA
 TAATTTGGAACACCATGTGTGCGCCCAACGTGTATGACCAATTCCTAACGTA
 CCATCCTCATCACTATTATCTGCAAC

Sequence 3413

step.1002g04.cons.ok

10 ACTGGTCCACACGTACACTTCCAAGTTGAAAGAGGTCGCCATGATGACATC
 ACAAACAGAGGGACAGTAACCCCTGCTAAATGGCTCAAAGGTCACGGTGGT
 GGAAAAGTTGGTGGTAGTGGTTCTGTAAACGCACGTAGAGCAATTCAAAGA
 GCACAATCTATTTTAGGTGGACGTTATAAATCGTCTTATATTACCGAACAA
 ATGATGAGAGTTGCCAAACGTGAGTCTAAC TTCCAATCAGATGCGGTTAAT
 15 AACTGGGACATCAACGCACAAAAAGGAACGCCTTCTAAAGGTATGTTCCAA
 ATGATTGAACCATCTTTTAGAGCATATGCTAAACCAGGACACGGAAACATC
 TTAAATCCAATGACGAAGCTATATCTGCTATGCGTTACATTGTAGGTAAG
 TGGGTTCCCTATTATGGGGAGTTGGAGAAGTGCATTTAAACGTGCTGGAGAT
 TATGCTTATGCTACAGGCGGGGTATTAAACACTGCTGGATTATATAATTTG
 20 GCAGAAGATGGATACCCTGAGATAGTAATCCCTACAGATCCAAGCAGACAA
 TCAGATGCGATGAAATTGTTACATCTTGCTGCAAGTAAATTAGTGGAAAT
 AACAGAAATAAACGACCTAACCAATTACGTACACCTAATGTTACTAGTAAT
 ACAGTTGATAATGCAGAATTACTACTACAAATGATAGAAAATCAACAGAAA
 CAAATAAACGTGTTAATGGAAATAGCACGAAGTAATAAACTATTGAAAAA
 25 CAACCGAAAGGTTTTTTCAGAACGCGATGTAAGTCAGGCACAAGGTTCAAGG
 TTAAGACTCGCTGCTTATAGCCAGGGAGGTTTTATAAATTGGAAAAATAAAA
 AGTAAAAATATTTAACGATCATTTTCAAGAAACACTAACGGATATTCCTCA
 TCTTAAGTTTCACAGAGGATAAATATCACATACAGATTTTGCTATACGCAA
 CATTGTTAAAACATGATAACTCAATCCACTCGCGAAATTATGATGATGAGA
 30 ACTAGCTGCTGGATAAGTGAAAAACGATCTTGATACCTTTTAAATTAAATG
 TCTAGTAATGCGTTGTAAGTTAGCATTTTCAATATCTAACATAAAATGCGA
 TAATTCCTCTTGTATTTCTGTAGGTGACATTTGGCGCACCGTCAACAAAGTC
 TTTAGTTGACACTTTTGCTTCAGCTTGTGCAAGACGATTTTGATGTATTTT
 CATCTGTTTGCGTCCACGATAATTGATGATATCACCTTTGACATGAACATAT
 35 TGTTTCTCGGTTTTAAAGTTTGCATATCATCTTTTCGTAGCAGTCCATAATTT
 AGCTTCAATATCACCACCTTTATCTTGCAGAAATAGTGTATGTAATCTTT
 ACCCTGAGCTGTAACACCTTGAGTTGCACGATGGATCAAGAAAAATGATC
 CACTGAGTCGCCGGGATTAAATTTTCTACATTTCTCATTTTTTCCCGCCT
 TCCTGTAATTTATTTAAAGTTAATACTTCTTTAGCCGGTATGACATGATCT
 40 TTTGTACAAGTAAAATAAAGTATTTGATAATGTTCTGATAGTTCTCTCAA
 TATTTCAACATACGTTCTTTACGATATTTATCAAAATGAACAAATGCATCA
 TCTACAACTCACTGGGAATGGATAATACGGTTTTAATACTTTAATAAGACTA
 ATACGTAAAGCCACATATAATAATCTTTTGTAGATTGACTCAACTCAACT
 GGCTCAAACTACTTGTCGGTTAGAATGCTTTACATGTATTTTATGATTTTCA
 45 GTATAATGAATCATATTGTAAGTACCATTTGTAAATTTTTAAAAATAGAT
 ACAGCTTCATTAATCACTTGTGGTAGACGCTTATCTTTTATTTGCTTGATA
 TGTTCTTCCACTAAAGCTTGCAATATAACTTAAGCTTGCCCAATCCTTAGCA
 ATATCGTTAAGTCTATTTTAAAGCTATAATATTCATGTCTTAATGTGCT
 AAAGTTCTATCTGTTTCCATATGATTAAATCTGAGCATTTAAATCACTAACT
 50 TCTGCTTGCAATTTCTAAAAATTGATCATTTGTAATCGTCAACTTGTTAGCC
 AATCTATGATCTTCTTCTTAGTTGAGCAGTAGTTTCTCACTTAATTGC
 GAACCTAATTCATAAGTGAATTTTGATTTTCTAAATATTGATTAAATCA
 TTAAATCGATTTAAATCACTTTGATATGTTTCAAAATGATGATGATGTGTA
 TAATAATCTTCTTCAATATCTACATCAACATAATCAATAATTGCTTAATC
 55 ATTTGATTGCTTTCACTTAAGCGTGAATTTCAATTGCTTTAATTCATTATTG
 AGTAATTCGCTTTCAGTTTGATTTTTATTCCAAGCCTCATTTTGTCTTCT
 GCATTTTTTAACCACGTGTTTACATCATGGAAGAAAGACATCTGATTAAAT
 AGTTTTAAATGTGGTTCAAGTTACAGATTGTGCACGTTCAAAAAGTGATTG
 ATATCTTTTAGCAGATTATTGCGTTGTTGATTAAATCAATAATGTATTTA

TCATGCGCTTTAATTTGACCAATGGTTGAGATACTTTCAACCAC TAATTCA
TCAGAAAGTTTTCTGATAAACGTAACCTCTTTTATTTC AATAATGCTA
TGTTTTAACTATCTAATCGCTCATTTGCAGTCGTTAATGATTGATTGATA
TATTGATGTTTTCTTCAAGTATTTTTTATTTTTTTTAGCATGACGCCAT
5 TGTTACGAACTTGTTGTTGATATCTAAATCGAAGTCAAGATTATATCT
TTTTCAAGTTGCGTGAGTTGGTTTTCTAAATCATTAATTTCTGACTTATT
GCTGTGCTATAATCTACTGCTTTAGATCTAGAAAAAATGATACCTACTACA
AAAATCACAGTTAATAGAGCAAAATTATACCAAAAAATAAGATTTGCAGTG
AAAAAAGAAAAATAGAAAGTGCTGCTGATAGAATAGTCAAAACAATAAAT
10 CCTATTCTCAAAAACCTTTTGTCTTTTATTTTGTGTTTGTCTTCTTCA
AAAGTTTCTTTTAACTTTTCATACAAGTCTCTTTTTCATGTAATTCTAAA
ACTTGTGTGATATTCCTTTTCTTTTCAAAGGTTTCGTGAGGAACAAGC
TCATTCTCAACTTGATTAATCTCATTAGAATTAGAGTTTCTTTTCA

15

Sequence 3414

step.1002g05.cons.ok

ACTTTGTTGATTC AACAGTCTCTGTAGCTAAAGAAGTTCAAAGTCCGACG
ACAATCTAGGTATTGCTATAGATGCCTATGGTGCAGGTAGTTTTATTGTAG
20 CAACTAAAAATTAAGGGAATGATTGCGGCTGAGGTTTCAGATGAGCGTTTCAG
CTTACATGACACGTAGCCATAAATGCACGCATGATTACTATGGGAGCTG
AAATTGTTGGAGATACGCTTGCTAAGAATGTCGCAAAAGAATTTGTCAATG
GTCATTACGATGGTGACGTCATCAAATCCGCGTAGATATGTTAAATAAAA
TGTTGTTAATAGGAGGAATAAAAAATGAAAATTGCAATAGGTTGCGATCATAT
25 TGTTACTGATACAAAAATGGAAGTTTCACAACACTTAAAATCACAGGGACA
TGAAGTGATAGATGTTGGAACCTATGATTTTCACACGTACACATTATCCGAT
TTATGGAAAAAGGTAGGAGAAAAAGTTGCGAGTGGTGAAGCAGATTTAGG
TGTATGTATTTGTGGTACTGGTGTAGGAATTAGTAATGCTGCAACAAAAGT
ACCAGGTGTTAGAACTGCTTTAGTTAGAGATATGACATCAGCGCTTTATTC
30 TAAAGAAGAGTTAAACGCCAATGTGTAAAGTTTGTGCGGTAAAGTAGCAGG
TGAATTATTTATTTTCGACATCGTTGATGCATTCATTGAGGCAGAGTACAA
ACCTACTGAAGAAAAATAAAAAATTAATTGCTAAAAATCAATCATTTAGAAGC
ACATAACAATGACCAAGCTGATCCACATTTCTTCGACGAGTTCTTAGAAAA
ATGGAATAAAGGTGAATATCACGATTAAGGAAGATGTCCAATGATTTTAAAC
35 ATTAACCTTTAAATCCCTCAGTTGATATATCTTATCCTTTAGATCAGTTTAA
TTTAGATACTGTTAATAGGGTATCTCAAACAAGCAAAACAGCAGGCGGTAA
AGGATTAATATGTTTACTAGAGTATGTCTGAATTTGGTGAAGACGTAATAGC
AAGTGGTTTTTTAGGTGGAGCATTAGGTCAATATATTGAAGAACAATAGA
AACTACTCGTATTAACAAGCATTTTTCAAAATCAAAGGCGAAACACGAAA
40 TTGTATTGCAATACTACATGAAGGACAACAACTGAAATCCTTGAAAAGGG
TCCCACGATTGAACTTAAGGAATCAGAGGAATTCAGTCAATCTATTAAA
ACTTTTCAAAGAACTGATGTGGCTGTTATGTCTGGTAGTCTGCCCAAAGG
ACTTAATACATGATTATTATGCGGATATTGTGAGATTAGCAAAAGAACAAGG
AATTTTGACCATTTTAGATAGCTCTGGTCAATCACTTGAGGAAGTCTTAT
45 TAGTAATGTGAAACCTACAGTAATTAACCCAATATAGATGAATTATCACA
ACTTTTAAATTAACAAAGTAACCAATGATATTAAAGAATTGAAAGCGGCAGT
AAGTCAGCCAATATTTAATGATATTGAATGGATTATTGTTTCATTGGGCAG
TGAAGGTGCTTTTGCAAAACATAATCAAAAATTTTATAAGGTGAATATTCC
CAACATTAAGTAGTTAATCCTGTTGGGTGAGGAGATCCACTGTAGCAGG
50 AATTGCTTCTGGACTATTTCATCAACAAACCGATGAAGAGTTATTAAAAAA
AGCAAATGCATTCCGAATGCTAAATGCAATGGAACAACAAACAGGTCATAT
TAATACAGATAAATTTGACGAAATATTCAAACAAATAGAAGTTATAGAGGT
GTAATTTATGACAAAATCACAAACAAAAGTGTCAATGAGAAATTAAAG
TAATCAAGAAGGTATTATTTTCAGCTTTAGCATTTGATCAACGTGGTGCATT
55 AAAAAGAATGATGGCAGAACATCAATCTGAAACACCAACAGTTGAACAAAT
AGAACAATTAAGTACTTGTTCGAAGAATTAACCAATATGCGTCTTC
AATTTTATTAGATCCAGAATATGGTTTACCAGCATCAGATGCTCGAAATAA
TGACTGCGGACTATTACTTGCATACGAAAAAACTGGATATGATGTGAATGC
GAAAGGTGCTTTGCCAGATTGCTTGGTAGAATGGTCTGCGAAACGTTTGAA

AGAGCAAGGGGCAATGCAGTTAAATTTTACTTTATTATGATGTAGATGA
CACAGAAGAAATTAACATACAAAAGAAAGCATATATTGAACGAATTGGTTC
AGAATGTGTTGCCGAAGATATTCCTTTCTTCTTGAAGTTTTAAACATATGA
CGACAAATATTCTGACAAATAAAAGTGCAGAATTCGCTAAAGTTAAGCCACG
5 TAAAGTTAATGAAGCAATGAAGTTATTCTCTGAAGATCGTTTTAATGTGGA
TGTACTTAAAGTTGAAGTACCTGTGAATATGAATTTTGTGGAAGGATTTTC
AGAAGGAGAAGTTGTTTATACTAAAGAAGAAGCTGCACAACATTTCCGTGA
TCAAGATGCAGCTACTCACTTACCATATATTTATTTAAGTGCAGGTGTATC
AGCAGAATTGTTCCAAGATACATTAAAAATTTGCGCATGATTCTGGTGCGCA
10 ATTCAATGGTGTTTTATGTGACGTGCCACATGGTCAGGAGCAGTTAAGGT
ATACATTTGAAGAAGGAGAGCAAGCTGCCAGAGAATGGTTGCGTACGGTAGG
ATTTAAGAATATTGATGATTTGAATACAGTATTGAAAACAACAGCTACATC
ATGGAAAAACAAATAATGTAAGGGAGGATATTCAAATGAATAGAGATGAGG
TACAAATTAATCGGATTTGAAATTTGTTGCTATGCTGGGGATGCACGTTCAA
15 AATTATTAGAAGCTTTAAATGCTGCTAAAGATAGTGAATTTGATAAAGCAG
AACAACCTGTAGAGGAAGCGAATGAATGTATTGCTAATGCACATAAAGCAC
AAACCAATCTTCTAGCTCAAGAGGCTAAAGGCGAGGATATCGCATATAGTA
TCACTATGATTTCATGGTCAAGACCATTTAATGACAACATTACTTTTAAAAG
ATTTAATGAAGCATTTAATTGAATTATACAAAAAGGGAGCTGACCCCTAT
20 GAATAAATTAATAGCATGGATAGAAAAAGGAAAGCCATTCTTTGAAAAAAT
ATCACGAAATATTTATTTAAGAGCGATTTCGTGATGGATTTATGTGCTGCTAT
TCCAATTATCTTATCTCAAGTATATTTATTTAATTACCTATGTACCAA
TGTGTTTGGTTTACTTGGAGTAAAACTATGGAAGGTATATTGATGAAACC
CTATAACTATAC
25

Sequence 3415

step.1002g07.cons.ok

TTTCATCTCAATATCCATACCAAGTAGCTCTTCTATTAAATCTTCGTGCG
30 ATACGATAGCATCCGTACCTCCAAATTCATCTAAAACAATCGCTAAATGTT
TTCGTGAAAACGGTCATTTTACGCAATACCCATTCTGCCCTATTATGTTTCAT
TTACAAATAAAGGGCTTGATGCATAATTAGTAATTGCGTCTCTTTATTTT
TACTCCAAGCTAATAAATATTTTGAGTGGAATACGCCGATGATATCATCTA
TATTTTCATCATATACTGGATATCTTGTATATGGGTTATTTCATCACTGTAT
35 CATAAGCTTCGTCATATGTTACTTCTTTGAAAAAGCTACAACATTAATAC
GAGGCGTGGTATCAACATCCTTAACCTTCAATTGTTCAAAGTCCATAACGT
TTTGAAGTCGAGTATTTTCTATCTCATTAATGACCTTCTCTACCCGCAA
TATTTAATAATGTACGAATTTCTTCTTTAGAAAACTTTTTTCAACAGGTT
GGCCTCGAGATAGTAAATGATTAATACCATCTGTCATCTTGTTTAATAAGA
40 TTGTAATGGGCTTGAGTACAATAACACATATATGAATGATAGGATACACAA
GTTTTGAAATTTTATCAGGAAATGTTGCAGCAATTGATTTAGGAATCACTT
CAGAAATAAGTATAATCACAATTGTTACAATTGCTGATGCTATTCCCACAT
TCACTCCAATATCAATCGCTAAGATTGTCACGAGTGTAGGTAAAAATAATAT
TCGCTACATTGTTACCGATTAATATAGTAGTGATAAATTCACTTGGTTTAT
45 CTAATAACTTTGTAAACCTTGCGCCTTACGATCTCCCTTTTTAGCTTCAG
TTTGAAACTTAGTGCGATTGCGGGCTGTTAATGCTGTCTACTACCTGAAA
AGAAAAATGAAACAAATATTAATAGTATAATTGCAATAATCATGATGCATT
TCTCCTTAAGATTTAGTCCGTATTAATACATTATTTCCCGAGAAATGAATT
ATCAAACCTTATCATTTATATTGAATTAACACAACACACGAATAAGAATGAA
50 CTAATGCAAAATATTTTATAGTTTGCCATAAATTAATAAAGAAGATATCG
ATTTAATTATTAATAGAAAGTTTAGAACAAAATGTATTTCTACTAAGTGAAT
ACATTCAAAAAGGTAATAAGAATAAAGCTATACAATTAGTCAATGATTTAA
TCATTATGAAAGAAGAACCATAAAAATTACTGGCATTAATTACTAGTAATT
ATAGATTATATTATCAATGTAAAATTCTTAGTCAAAAAGGCTATAGTGGGC
55 AACAAATTTGCAAAGACTGTAAATGCACACCCTTATAGAGTAAACCTAGCAC
TCAATCAATCCCACATTTATAAAGTAAAGTTTGTAAACATCATAAACG
CTTGTGCGAGAGACTGACTACAAACTAAAATCATCTTATATGGATAAACAAC
TCATTTTGAATTTATTCATCTATCTCTATAAAAAATTTGAGGTAGGAGAT
ATAAAATAAATCCTATCGACAATAGAAATATTGGTTGTAAAGTAATAAAAA

AGAGCCTAGGACAATTCGTTTTGTCCCTAAGCTCTGTTTTAATAATCATTAT
 TTATTAGCTGACATAAGTTTAGATTTAATACGGTCAGCTTTATTAGAATGG
 ATTAATTTACTTTTGAGATGCTTTGTCAACTTGTGATAGCAAATCTTAAT
 AATTCATCTTTGTTTTTCAGCATCAGTAGAGATAGCTGTTTTAGCTCGTTTC
 5 ACAGCTGTACGCATAGCGTTTTCTTTAGAAATATTTCTTTCTTCAGCTGTT
 TCAGTTGTTCTTACACGTTTGATTGCAGATTTAATATTTGGCATTTACTGTC
 ACCTCCTAAAAGTGATCATAACTTATCAAAATTTATTTGATTACAACAAGA
 AATATTTTATCAAAATGCTTATATTTGTGCAATCATTTATTTAAAAGTAAC
 TTCATTTGAATTTATTACTTTTAAATTGTTGGAATGAACTACTGAACTAGA
 10 TATAATAGTTATGTTAGAATTTACATTTAATTAATTATGTCAAAGATATAT
 TGCATTAAAGTAATAAAAACGTTACTATATCAAAGATACATAGAAGTGACA
 GGTTATAAAGATGAAAGCGAGAAGGATAAAATGGATAAGCAAGAACGATAC
 AATAGAAGAGAAAATATTAGAAATTTCTCCATTATTGCTCATATAGACCAT
 GGTAAATCGACATTAGCTGATCGAATTTTAGAGAAATACAAATCAGTTGAA
 15 ACTCGAGAAATGCAAGATCAATTTACTTGACTCTATGGATTTGGAAAGAGAA
 CGAGGCATCACTATTAACTAAATGCTGTTTCGATTAAAATACGAAGCTAAA
 GATGGAGAACTTACACATTTTCAATTTGATAGATACACCAGGACATGTCGAC
 TTTACATATGAGGTTTTCTCGCTCATTAGCTGCATGTGAAGGTGCAATTTCTT
 GTAGTTGATGCTGCCCAAGGTATAGAAGCACAAACCTTAGCAAACGTTTAT
 20 TTAGCATTTAGATAACGATTTTGGAACTTTTGCCAGTTGTTAATAAAATAGAC
 TTGCCTGCAGCTGAGCCCGATAGAGTTAAGCAAGAATTAGAAGATGTTATA
 GGTATAGATCAAGAAGATGTAGTACTTGCAAGTGCTAAGTCAAATATAGGT
 ATTTGAAGAAATTTTAGAGAAAATAGTTGATGTTGTACCAGCACCGGACGGT
 GATCCAGAAGCCCCACTTAAAGCACTTATCTTTGATTCAGAATATGATCCA
 25 TACAGAGGAGTAATATCTTCAATTTCGAATTATTGATGGTGTTGTTAAAGCT
 GGAGATAGGATTAATAATGATGGCTACCGGTAAAGAAATTTGAAGTTACAGAA
 GTCGGAATCAATACCCCTAAGCAACTACCGGTAGAAGAATTAACAGTTGGT
 GATGTGGGTTATATTATCGCAAGTATCAAAAATGTTGATGATTCTAGAGTA
 GGTGACACAATTTACTTTAGCTGAAAAGACCTGCTGACAAACCGTTACAAGGA
 30 TATAAAAAGATGAATCCAATGGTATTTTGTGGTCTATTCCTTATTGACAAT
 AAAGACTATAATGACCTAAGAGAAGCTTTAGAAAAATTACAA

Sequence 3416

35 step.1002g10.cons.ok
 TGTAATACCAATAAGTTTCAGTTCCCTTGAATGAGAAGCCTGCGACTAGAAA
 TACACCAAGAATTGATAGCAAGCTACCTCCTAAATTTCCACCTAGTATTGG
 CCCATCTCCTTTTGTAAACGTATCAAAATCCTACAAATTTCTCCACCCATAAT
 CCCTAAAATAGTTAAGATACCAATTCCTATAAATATGATGACTGTAACCTAC
 40 TTTGATAAGTGCGAACCAATATTCACCTCTCTCCATATACTCTAACGGATAA
 AGAATTAAGCGCGAAAATAATAAAGGAAAATACAACTCCAGACCCCAAGC
 TGGTATACCTTGATAGGGGACCAATATTGTATAAATTTGCGCTGCAATAGT
 AACATCAGCTGCTACTGTTATCACCCAGTTGAACCAATAATTCATCCTTAA
 AGCAAAACCTAGTGAAGGATCGACAAAGCGTGTAGCATAAGTACTAAATGA
 45 ACCTGACACAGGCAAAATATGTGCGCCATCTCCCCCTAACGACGTCATTAGAAA
 GAACACCATCGCTCCAATAACTGCATAGGCAATCAAAGCACCCAATGCACC
 TGCGTCATGTATAGCTCCACCAGAAGTCATGAATAGACCAGTTCCCTATACA
 TCCCCCAATGGCAATCATTGAAATATGTCTATCTTTAATCCTCTTTGAAC
 ACCGTCATTTGTTTATTATCTTTTCATAGAAGCTCCCTTTCCCTTCATA
 50 AGTAGAGGCTAACTTGTACCACATCAATGCATATCATTGTTGTAGATAAAA
 TTTGGTATAGATGATATAAAAGTAAAATAATATTAATATCATTTTACTTTT
 ATATCACCTGTCTTTTATATCTGTATCAATAATGTAAAAAGTGGTTACATAC
 CTTTAAGATATGCAACCCTAATGTTTATCTATATAAGGTAGCACATCACA
 TAAAGTGACAGTACAACCTTCTATTCGAAAGTTGTCCCAACAATGCTAATTT
 55 ATGGAATAACATCATTTCGGCATTTATTACCTTTCACTTATTCAACACATG
 TCCCATACGACTTCTGATAAGTTAGTCATTAATAATGCAACCTCTAACTCT
 ATTTATTATATTAATTTGTAATATACATGATATACCTTTATATTTTCAAGTT
 TATCTATGCTTTTTCAAACGGATTTTCATCAATAACATTCCAAATAAATTCA
 TCTTTTCAACTGATTCTTGTTTTTTCAGAACCATATTTTCTTACATTTACT

TCTTGATTCTCTACTTCTTGGTCACCTACAACAATCTGATAAGGTATTTTT
 TTCATTTGAGCTTCACGAATTTATATCCCATTTTTTCATTACGGTCATCA
 ATTTCAACCGGACACCTTGGGATTTTAGTTTCATCTTGTAAGTCTTGCA
 TAATCATAATGTAAATCTATATTTACAGGAATAATTTCAACTTGCATAGGC
 5 GCCAACCAAGTTGGAAATGCACCTTTTGTCTTCTGTTAAAAATGCAACA
 AAACGTTCCATAGTAGAACTACACCACGGTGTATAACTACAGGACGATGT
 TGTTCTCCATCTTGACCAATGTACGTTAAGTCAAAACGTTCTGGTAAAAGA
 AAATCAAGTTGTGCTGTTGATAGAGTTTCTTCTTTTCCCATAGCTGTTTT
 ACTTGAACATCTAACTTAGGTCCATAGAATGCTGCCTCACCAATAGCTTCT
 10 TCATAAGTTAAACCTAATTCATCTGATGCTTCTTAAGCATGGATTCAGCT
 TTTTCCCACATTTTCATCATCATCAAAGTACTTATGCTTATCTTCAGGATCT
 CTATAACTCAATCTGAAGCGATAATCTTCAAAACCAAAATCTTTGTACACA
 TCTTGAATCATATTAACCTACACGTTTAAATTTCTTTTAATTTGATCAGGT
 CTAACGAAAAATGCGGAATCATTCAATGTCATTCCTCGAACACGTTGTA
 15 CCTGATACGACCACTTGCCTCGTAACGATGCATAGTAACCAATTCAGCA
 ATACGTATAGGTAATTCGCGATAAGAATGAGGTTTGTTTTTATAAATCATC
 ATATGATGTGGACAGTTTCATTTGGTCTTAAGACCATTGCTTCGTCCTCATCT
 AACTTCATTTGCTGGGAACATATCTTCTTGATAATGATCCAGTGACAGAT
 GTTTTATATAAATCAACATTTGGCTAATACTGGTGTGTAAACATGATCGTAT
 20 CCCATACTTACTTCTTTATCGACAATAAATGCTTCTATTTCCCTACGTATT
 GTAGCACCATTGGTAACCATAATGGTAACCAGCACCAACGAGTTGATTG
 TTTGTAAACAATTTCAATCTTTACCAATTTTACGATGATCAGCTCACGA
 CGTTCTTCCAACATTTTTTAGATGTGCTTTCAAATCTTTTTTGTCAAAGAAT
 GCTGTACCATAAATTCGTTGTAACATTTTATTATCACTATTTCCACGCCAA
 25 TAAGCACCAGCTGTAGATAAATAGTTTGAACCTTTAATTTTAGAAGTAGAA
 GGTACGTGTACACCTCGACATAAATCAGTAAATTCACCTTGAGTATAAAGT
 GTTACACTCTCATCTTCAGGAATTCATCAATAAGTTCTAATTTATAAGGG
 TCATCCTTGAAGAAGTCTTTTGTCTTTTCTTTACTAACTACTTCTTACA
 ATTTTATGATTTTTCGTTTCAATTTGTTTCTCAATTTTATCA
 30 AAATCATCCGATGAAACCTTATCATCCATATCAAAATCATAATAGAATCCG
 CCTTCTATTACAGGTCCAACCTCAAATTTAAGCTCTCCGTATAACGTTTT
 AATGCTTTGCCATTAAATGAGCTGTTGAATGACGAAGTACTTCTAACGCT
 TCTTCACTCCAGGAGTAATAATTTCAATAGCTCCATCTTGTCTTAAAGGG
 CGTGTAAATCTACAAGTTGACCATTGAATTTTCCCGCAACTGCTTTTTTT
 35 CTTAATCCTGGACTAATTGATTGAGCGATGCTTCTGTAGTAGT

Sequence 3417

step.1002g12.cons.ok

40 ATACTGCGATATTAACCGATGAAGAAATACGTGAAAAAACAAATCTTTTC
 AAAAAGAGTTAGCAGAAATGAAGATGTAAAAAACAAATGATTATTTAG
 ATAAAAATTTACCTGAAGCTTATGCACTTGTACGTGAGGGTCAAAGAGAG
 TATTTAATATGATTCTTATAAAGTACAAGTAATGGGTGGTATTGCTATAC
 ATAAAGGTGATATTGCAGAAATGAGAACAGGTGAAGGAAAACATTGACTG
 45 CAACCATGCCGACGTATTGAAATGCTTTAGCTGGTAGAGGTGTACATGTTA
 TTACAGTCAATGAATATCTATCAAGTTCACAAAGTGAAGAAATGGCTGAAC
 TATATAACTATCTTGGCTTAACTGTAGGTTTGAACCTAAATAGTAAGTCAA
 CTGAAGAAAAACGTGAGGCTTACGCACAAGATATCACTTATAGTACGAATA
 ATGAACCTGGGTTTGATTATCTTAGAGATAATATGGTGAACATGCTGAAG
 50 AGAGAGTAATGCGTCTCTACATTTTGCAATTTATGATGAGGTCGATTCCA
 TATTGATCGACGAAGCAAGAACACCTTTAATTATTTCTGGTGAAGCGGAAA
 AATCTACTTCTTTATATACTCAAGCAAATGTTTTTGCAAAATGCTTAAAG
 CGGAAGATGATTATAATTATGATGAAAAAACCAAAGCTGTACATCTTACAG
 AACAAAGGTGCAGATAAAGCTGAACGTATGTTCAAAGTAGATAATCTTTATG
 55 ATGTTCAAATGTGGAAGTGATTAGTCATATTAATACAGCTTTAAGAGCTC
 ATGTTACTTTGCAACGCGATGTTGATTACATGGTCGTTGACGGTGAAGTAT
 TAATTGTTGACCAATTTACTGGACGTACAATGCCTGGACGTCGTTTTCTG
 AAGGTTTACACCAAGCAATTGAGGCTAAAGAAGGTGTAGCAATTCAAAATG
 AGTCTAAACGATGGCATCCATTACTTTCCAAACTATTTCAGAAATGTATA

ATAAGTTAGCGGGGATGACTGGTACAGCGAAAACCGAAGAGGAAGAATTTTC
GTAATATCTATAATATGACAGTTACCCAAATTTCCAACAAACAAACCTGTTTC
AACGTAAAGATAATTCAGACTTAATTTATATTAGTCAAAAAGGAAAGTTTG
ATGCGGTAGTTGAAGATGTTGTAGAAAAACATAAAAAAGGACAACCCGTCT
5 TACTAGGTACTGTTGCTGTTGAGACTTCTGAATATATTTCAAATTTACTAA
AAAAACGTGGTGTGACACATGACGTATTAAACGCTAAAAATCATGAACGCG
AAGCTGAAATCGTTTCAAACGCGGGGCAAAAAGGTGCAGTTACAATTGCCA
CAAATATGGCTGGACGTGGAACAGATATTAAACTTTGGTGATGGTGTGAAG
AGTTAGGTGGACTTGCTGTTATTGGTACTGAGCGTCATGAATCAAGACGTA
10 TTGATGATCAATTACGTGGACGTTTCAGGACGCCAAGGTGATAGAGGAGATA
GTCGTTTTTACCTATCTTTACAAGATGAATTAATGGTACGTTTTGGTTCAG
AACGCTTACAGAAAATGATGAACCGTTTAGGAATGGATGATTCAACGCCAA
TCGAGTCGAAAATGGTATCTCGAGCTGTAGAATCAGCTCAAAAACGAGTAG
AAGGTAATAACTTTGACGCGCGTAAACGTATTCTAGAATACGATGAAGTTT
15 TACGTAAAGCAACGTGAAATTATTTATAATGAGCGTAATGAAATCATTGATA
GTGAAGAAAGTTCTCAAGTCGTTAACGCGATGTTACGTTCTACATTGCAAC
GTGCGATTAATCATTTTTATTAATGAAGAAGACGATAATCCTGACTACACGC
CATTTATCAATTACGTTAATGATGTGTTCTTGTGAATTATTAATGTTAAA
TGCTATCCATGTGCCATTACACACCGGTAAAAATTGAAGAAATGACACGTGT
20 ACTTCGTGAAAAATAAGCATAATATATTTAAGCGTCTGGATTATCATCCCC
AGACGCTTTTAAATGTTTAGAGAATTATCAAATTTTATTTGAAATGCTTTC
GCTAAAACAAAGATAAACATATCTTTTCGAGGAGTTTATCTAGTATACGTA
CATTTCTCCCTCAACAAGCTAAAAACATAACTATCTGAGAATTTCCCTTGT
AATCGTTTCATTACTTCTCAATATACCTTTGCGTGAACCCAAGTCTAATTGG
25 TATAGCTCTGCTTTTCTTATTTTAGTAGACATTCGTATTTCTATACGATT
AATATCATACACTTCATATACATAGCGAATTAACGCTTTAGTACATTTAGT
CATAATACCTTTCTTTTGAAGTCTTCAGCTAAATAATAACCAATTGAAGT
TGTTTTTATTAACATAATCTAAGTAATGCAATCCTATGACTCCAATCAATTC
TTTATTTACTCCATATTTCCACAATGAAATCCATTACCATCGATAAAATTGTTG
30 CAACGCCGAATGGATAAAGTGTTTACTATCTTCAACTTTCTTCGTATGTTTC
AACAAAAGGCAGAAATTCAGCTAAATAGTCACGATTGCTATCTACTAATTT
AAATAACTGTTTCGGCTTCTCGTTCTTCTAATATTTTAAATTTAATTTTTT
ATTTATTTGATAACTAAACATACAAAATCATCCTTTCTTAGAATTTCTGT
CATTTTAGTATACGCGAGTATAGGATTATGTCATAAACTTGATTACTAGA
35 ATTCGATATTCTTTAAGTTACAAGTGTGTATGTAGTATTTCTCTTGCTTCT
ATTGTCCTTACTTTACAACCAATATTAGAGTGAAAAAGCTTAACTTAAGAT
AGATAATCATCAACAACACAAACGTTAAAAACACTGAGAAAACTTTTTTAA
AATATACTTCTCAATATTGATTAGTAAATCCAAAATATGAATAATAAGTA
GACATTAAACAATCTTTATTTTATTCTATAAAATAATAAAATAAATTATT
40 AGATTAATACATTATTTTTTTATTTCTAATTTAAGTTGAAATCGTATGTAA
TCGGAGGAAACCATAATGATATATTCTATCGGTAAGAATTTAGGTAATAAAA
TTAACAGGTATAGAACAAGCTATGATCAATAGATTAAAGCTATTTAAAGAT
AATTTAGTCCCAATAAATCATATTACATCTTTGGTCACCACGTTTATAT
ATGCATGCACATTGTTAAACATCGATTCAAAAGATATTTTCAGTCTTTAC
45 GATTTTCTACAAGATAGTATTAACCTTTGAGAAAAACATATTGATTGGATA
AATTAATTGGCAAAATATATGTAATTATACCTTAAAAATTCGTTGAAAAACG
AATGATATTTAAATATACGATAACGACACATATAAAATGTATGTGCATTTT
GTTGATTCAAATTATCAAACCTTTAGACTATATTAACCATTTTGATATACAA
CAACGTAAAAATTCGAAGAGATTTTACGATACAAGAGGCTTTTAAAGTTGT
50 AGTAGAATTTTAACCTCTCAACAAAAAGTCGTGATGGAACAATTTTACAC
CCTACACAAAAAGTTAAATTTCAAAAATATTACAACCCTGAGCACGAACAT
CCTACGGTACAATCTATCATTT

55 Sequence 3418

step.1002h01.cons.ok

CTTGAAATATCATTTGAATTTTCTTATTAACCTTCAATCGCTCTTTACGTT
TTCTAATTTGATGAATATCTTTACCATCATACTCGATTGCGCCATCAGTAA
TATCATTTAATTTGATAATGGCTTTTCCAGTCGTTGATTTACCACATCCAG

ACTCTCCTACTAGGCCTAAAGTTTCCCCTTTATAAACATCAAAACTTATAT
 TATCAATCGCTTTAACTTCATTATGTTTACCTTGATTAATAACTGTTTAA
 AATTTTTAACTTGTAATAATACATTTTCATTACTCATTAAGCCACCCTC
 TCTACTCTATGTGGTTGTTTCATAATTACTTGGCATTGTTCTTAATCGTTTT
 5 TGAACCATGAAGGTGGCGTAACTTTTGGAGCTCTTGCATCTAATAACCAA
 GATTTAACAAAATGTGTGGGTGAAATTTTGAACCAAGGTGGTTCTTCTTA
 AAATCAATATCTAAAGCATATCGACTTCTACGTGCGAAAGCATCACCAATT
 GGTGGATGAAGTAAATCTGGTGGTGTACCTGGAATTGCAATTAAGTCCGTG
 TCATTACTGGTTGTTAAATCAGGCATTGAAGAAAGCAATCCCCAGGTATAG
 10 GGATGTTTAGGATCATAAATATTTTCATTACATCCCCTGTTTCAACCATC
 TGTCCCCATACATTACGGCTACTTTGTCCGCAATATTTGCTACAACGCCT
 AAATCATGCGTAATAAGATAATTGAAGTTTCAATCTTATTTGTAGTTCT
 TTCATAAGATCTAAATTTGAGCCTGCATTGTACATCTAAAGCAGTTGTA
 GGCTCATCAGCAATTAATATTTTAGGCTCACATGCCAATGCTATTGCAATA
 15 ACTATTCTCTGACGTTGTCTCCTGAAAATTGATGTGGATAAGCTTTAAAT
 CGTTTTTTCAGCAGGAGTAACCAACTAAATTCAAAATTTCCAATGCTCTT
 TGGTTGGCCTTTGCTTTTACTTTACTTTACTTTTCTTATGTTTAAATCAAAGGTTCATG
 ACTTGGCTTTCCGATTGTGCATTGTTGGATTAAAGAAGTCATAGGATCCTGA
 AATATCATTGAAATATCTCGACCTCTTAGCTGTATCAGTTCTTTTTCACTT
 20 TTCTGAGCTAAGTCCTCACCTAAAAATAAGATTTCTCCCTTTTTTATCTT
 CCTGTATCCTTTTGAAAAAGTTTGTAAATTGCCTTAGTTGTTACAGATTTT
 CCAGATCCAGATTCTCCAACATGGCTAACGTTTCCCCTTTATTTAAATGA
 AAATCCACGCCTCTGACAGCTTGCACTTCTCCTGCAGCAATATCAAAGGAA
 ACGTGCAAAATCATTTACTTCTTAATACCGTTTTCTGTCTACTTTTACCTCC
 25 TTTATTTTCGCATTTTGGGACGTTTAAACGTTTAGCTTTCTTTCTAATA
 CTTTTCGTTTTTACTTTGTTTAGCAGCGTCAAAGATTAAGCTGTATAACCA
 GATTTCGTTGATGATTACCATGTTTCGTTTTTGACCTGATGCACTAAGTTGG
 TGGCTCAGCTTATCTTCATCATCAACATGATTTCTAATTCGATTATCAGAT
 CTTGTGTAACTAATATTTCTGCCACGCTTTTACCTAAAAATATGGTTCT
 30 CCATCTACTTCTATTTTCTTACTGGTAAATCTTCAAAATTAATACTTGT
 AAATCTTGATATGTTTATGCTCCTTTCTGCTATACTCCTTATAAAAGGA
 GGTGAAAATATGAAAAGTTTTATTTATGCGTATGATTAAATAACCAAAAG
 GATTATCCAAAATTAATAGAGCGTATTGAGGATTATCCTAATGTTGCTAAA
 ATCAATAAATCAGTTTGGTTTATTAATTCAACTAATGATGCTAAAATATT
 35 AGAAACGAATTAATAATGTTTATGATAGCGATGATAGTTTGTTCGTTGGT
 AAGCTGACTGGTGAAGCCGCATGGTCTAATGTAATTTGCAGTTCACAAAT
 TTAAAAGATTATCTTTTAGTATCGTTAATCTTGAGCCTCTCTTTTGGAGGG
 CTTATTATAATTTTCGTCAAATTTCTTCTGCATTTTATATTAGCTGCGACT
 TCTTCTAACTTTCAAAAACATAATGTATATCCACTATCTTTGTTTCTTTT
 40 ACAAAAACCTGTATAATACCTTGTTTTTCAGTCATTGGAATTCCTCCTTAT
 AGTTAAAACCTTCTTTTGGCGTAAGTCTTGTAAAAAATATCTCTTCC
 TTCTTGAGGTGTTAAATCTAACGCGAAATAAATTCATTTATTACTGGATA
 TGAAGTTTAGTTCTTCCGTGTATCATATTGGACAATGTATCTCTATTAAAC
 45 ACCATTTCTTCAGAAAGAGTTTGTATGTTGTGTCTTTTAAAGCCATTTT
 TGATTTTAAAGTTTAGTGTGATAGGCATTTTCAATTCACCACCTTTTCGTA
 TTACGTAAGTAATCTTATCATGGCTGTACAAAATAGGTCAAGCATTTTACG
 AAAGTTTTTAAAGAAAAATATTGCAAAATCACGAAAGTTTCTCTTATAATAT
 AGTTATAAAGTAAAAGGAGCTGTATTACGATGTGCTTTTCAAAAAGAAATGA
 AACAATCAAGAGAAAAACAAGGTATGACTTTAGCTGAAGTGAAGAAAAA
 50 TCGGTAAACTGAAGCTACTGTACAACGTTATGAAAGCGGGAATATTAAAA
 ATCTTAAAAATGATACTATTGAAAGTATAGCTACTGCATTAAATGTTAACC
 CTGCTTTCTTGATGGGTGGATAGAAGAAGTTGAGGAACAACCACAACATC
 GTGCAGCGCATCTTGATGGTGATTTAACTGACGAAGATGGCAAGAAATTC
 TTGATTACGCTGAATACATAAGAAGTAAAAAGAAATAAAGGGTGTTTTATG
 55 TGGGGAATATGAAGATATGTTAATTGAACATGACTATATTGAAGTCATTG
 AATGTGATAACTTACCTAAAAGGTATCTGGTTTGTGGCTTGGAGATATGA
 TTTTAATTAATCGTAACCTTGCCTATTACTTCCAACTTGAAACACTTGCAG
 AGGAACCTCGCTATAACGAACCTTACATATGGAATATAGTTGATCAAAGTA
 GTTTTAATCATAGAAAATTTGAAGGTATGCACGTAGGTAGCCTATGAAA

AGTTAATCCCTCTTAAAGATATTGTAAAAGCATTTTTGCAAGGCATTCATG
ACTTGTATGAACTTGCTAATTTTGTGA

5 Sequence 3419
 step.1002h02.cons.ok
 AAATAAACGAGAAAACCATAATCATGATAATCAACCATACAAATAGTACAA
 ATGTATTAAAGGTTGGATACCTTATATCATATGTTTTAATGTTGAAACCTT
 TATTATAAATAAATGAATTGAAAAATGCTTGTGCTGACTGTTTCGAATCA
10 TAGTCAGTTGTAAGGGAGCGAGACGACGAATCTGAAAAGAATCGGTCTCGC
 TTTTTTAGTATAAAAAATTTAATTATTATTTTAAAAATGATAAAGTGATTCT
 ATAGTTAAATTTAGTGCAGGATACCTTAAAAAATATATTGAAAAAATGT
 CGCTATTACAAAATGTTCTAAAGTGACAAATTCATTTAAAAACAAAAAAG
 AAGTTTTAAAAATTTATATTTCTGTCTATATATGAGTTGTGCACGAAATGAC
15 AAACGCAATTCACATATGAAATATTTATCACTTTACAATAATATGTTGTTA
 AGTTGATGTAGATAAAATTCATTGTTTTACGTATAAGTCACTATCATTGTTT
 TTTAGGATTCATTTTAATTATATCCAATCGACTAAAGTGAAAAGGAGGAAC
 ATATTAAATGAATTGGATATCGATTATTTTATTTATTATGGTCGTATGTGG
 CATTTCTTTTATGCCTATTTGCAATCAAGAAAAATTTAAACGAATAGTTC
20 AGATGGTTATTTTATGGGAGGAAATAGCCTTACTGGATTACAGTTGCCTC
 TACAATTATCATGACCAATTTGTGACAGAACAAATGTTGGGCAAAATGG
 TCAAAGCTATGCACAAGGAATGGAAGTAATGGCTTGGGAGGTTACGGCTGC
 CGTAGCAGTTGTATTGTTGGCCTGGGTCTTCTTCCTAAATATCTAAAAATA
 CGGTGTTAATACAATTTCTGAGTTTTTGGAATTACGTTATGATACATTTAC
25 TAAACGTTTTGTCTCCATTTTATTTATCTTTACCTATGTCGTATCTTCTT
 GCCTGTAGTATTGTATTCTGGTTCCCTAGTGTTCATAAAATGTTTAAAGT
 TGATGAATACTTAGGTGTAAGTAGTTCAACTGCTGTCATTATTATTTTCATC
 TATTATTGGTATAATTGGCATTTATTTACTTATTTATAGGTGGTTTATCGTT
 AAGTGCTTTTAGTGATTTCGATTTATGGCATGGCTTTAATTATAGCTGGACT
30 TGCGATTACAATATTAGGTCTAGGTCAATTAGGAGATGGCAACTTCCTACA
 TGGTTTTGACAAAAATCGTGCAAGACACGCCTGAGAAATGAATGGTTTGG
 TAAGGTGGACTCGGATGTTGTACCTTGGCCAACCTATTCTTCGGTATGTT
 CTTTAACAATTTATCTTCTGGTGCGCAAACCAGATGATAGTTCAAAAAGC
 ACTCGCAGCTAAAAATTTAAAAGAATCTCAAAAAGGTGCAATATATTTAAG
35 CTTATTTAAAGTGTTCCGACCTTTATTTACAGTCTTACCAGGCGTAGTAGC
 ATTTAACTATTTTAATGGTAGTCTTGATAAATCAGATAACGCTTACCCTGC
 ACTTGTAACCTTCAGTATTACCAGAATGGGCATTCGGCTTATTTGGTGCGGT
 TATTTTTGGTGCAATATTGAGCTCATTTGTTGGCTCATTGAATAGTACAAC
 TACACTATTAACACTCGATTTCTATAAACCTATTTTGGAAAAAATAAATC
40 AGATAAACATATTGCTCGAGTAGGCCATATTGCTACTGTAGTCATTGGAGT
 TATTGTTGTAGCACTTGCACCAGTCATCTCATTTATCCCTAGTGGTCTTTA
 TGCAGTAGTTCAACAGTTTAATGGTGTGTATAGTATGCCAGTGCTAGCTTT
 AATTTTAATCGCTTTCTTTCTAAACGCACATCTAAATTGGGCGCTAAAGT
 GACACTCTTCACACATATAATTTTATACGCTATAATCAGCTTTGTATTTAC
45 AGAAATTAATTATCTATACACATTTAGTGTATTATCTTTGTAGATTTAAT
 TATTATTTTGATTTTTTAACAAAAGTTAACTATCTAGTGAGTTTGATTTAAG
 CACGCACCAACCGAAAGTAGATATGACGCCATGGAAATATCGTTACGTTGC
 AGGTATTTATGTTCTTGGCTTTAGTAGTAGTAAGTTATATTATCTTCTCACC
 ACTCGTGTGGCAAAATAAATGTTACAATTAAACTTAATGAGTAAGAAAA
50 TAAAAGGTACGTTAAGGACGTTTATCAGCGTGATCACGTGCACTTTAACT
 ATGAATAAAGTTGATTTAAAAAATTAGAAAGGAGGGACAACGACGATGCAAA
 CAGTCGAATTATACCTTCGCCAGGTATAGCACATCAACATGCAAAAAAAA
 TAATTCCAAATGTTAAACAGTTATTGTCAAAGCGTACTAAACATAGTCAAT
 GGAATTTTCGACATCAAAGTCGATCTCATGATAGGATCTGCAGAGGATGTAC
55 ATGAAAGTGTTGAAAAAGCAGCACAAATTAAGAGGAACATCAGTGGGATT
 ACGTTGTTTGTGATGACAGATTTGCCTAGTATTTAGATAATAAAGTGTTG
 TCAGCCGACTTTAATAGTGACAAACATGTTGCAATGCTATCATTTACCGTCAC
 TAGGTTTTATTGATTTGAAGCGCAAGCTAGTTAAACGATGACTTCATTGA
 TTGAACAATTATATTATAATCAACCGAAAGACAAAAATGCGCCACATCCTT

TTGTACGCGTGAAGGCTGTAGAACCTGACGAAGACGCCACATCAAAACAAC
 GATATATTAATATTTTATTTATCATAAGTTGGATTTCAGTTAATTGGTGGAC
 TGACACGAGCAAATCAGCCTTGAAAAACATCTTTAATTTAAGAAAAATCA
 TTTTCAGTTGCCCTTTGCAACAGGAACCTTATGTCTCAATATTTTCAATGCCAT
 5 GGGAAATTAAGCGTGATTTATTCACCGCTTCGACTTATCATATTGATGGTGA
 TTGCTATACCTGGGATGGCTGGATGGCTATTCTATGCGCATCAATTGATTG
 AAAAGAAAACCTGCTAAATCTCAGCGTGATATCGATATATTTATAAATCAA
 CCACACTTGTTCACCTAAGTTTGATTACACTCATAAATTATGTCATTTTAT
 ATTTATTTGTTAATCATCAGTATTACACTCTTTGTCCCTGTGGAATTATTTA
 10 ATAGTTGGACGAGTGCCCAATCACAATTTACGTTCTCAAATTATATGAGAT
 TGATTTGGTTTGTATCATCATTTAGGACTTTTAGCTGGAGCTATGGGATCAA
 CTGTTGAAAATGAAGAGAAAATACGTCGTATTACTTATTCTTATAGACAAT
 ATCATCGTTATAAAGAAGCTGGCAAGAACAAAAGAACAAGAACTTCTCG
 TGATGTATCAACAACAAATGTGCAACAACAACTTCAAGTAAAGATGAAAA
 15 TAATGAACAATATGAAGGTAAAAACAAGGACATAGAGAGGAGGATGACGC
 ATGACAAATCAAAAAACGTGGGTCTAGTCGTCGCTCCAGGTGTTACTGAA
 CGCCTTGCAAAAATCTCATACAAGAAATGCCTAAAATGTTATCTACGCAT
 TATGATCATCAGCAAGAATGGATTTTGTGATTTA

20

Sequence 3420

step.1002h05.cons.ok

CCAACGTGAGATAATCAGCACCTTTTTCATTAATAATAGCTGCGGCCGATT
 CTAATATTAATTGTCTTTTGGAAAGTTCTAGGCATTTCGTCACCTTCTTTAT
 25 CGAATTCATAAATAATATATAATTGTAATTATTTAACACTTTTTCACATCC
 ATACTTATCGAGTGATTTTATTTATAAACAACCTTTCTTAAATTCAACTT
 TATTAACATGAATTGTATTTATTAATAACGATTTACTTATTACTATATGTC
 ACAACATATAAAAAAGTAACTTATCATATAATATATGTTCCCTCTGTCAAC
 GTCGGTTAATTATGAAAAATATTAAGTAGCTCCATTCCGTTGCTTAATAC
 30 TTTTTTATTCACTTATAAAAAATCCAAGAAAAGAGAGCGGATTTGCCCCC
 TCTTTTTTGTGTAGTAGTACTATTAATTTTATCTCACTAACAATTAATAT
 ACTTATTAATAAAATGATGTATTTAAAAATTTATAATAATCTTGAAATATAA
 TTATTAAGAAATTCGCAGTTATATCATTTTATTAACCTCAGCCTTATTGTACG
 TGGGACGATTTTTTAGCAGTCATAAAGTATACAATAATACCAATAATGAAA
 35 CCAATAAGTGCTGGAACAATCCATCCCATACCAATGTTAAAGAATGGGAGA
 TAATTTTCACCAAATCCAATGATTGTTGTGCAAAATTTGTATTAAACAAAG
 AACTCTGGACTTGCTTTTACACCATCTACAATGCCGCCACCATTGTAAAT
 AATGTTGTAAATCGATAGACAATTTTGAATGTTTAAATAATGGACTAAGT
 AATGCTAATAAAATTAAGTAATCGCTAAAGGATAAATGAACATCAACACT
 40 GGTGTGAATACATAATAATTTTGTAAACCTACATTAGCAAATATACAA
 GCTAATATACTCACCCAGTAGCTAACCAGGATAGTTAGATTTAGGGAAT
 AACTCTGTAAATGTTTCCGAAAAGGCTGTGATCAATCCTATTGCTGTTTTT
 AGACATGCCACAATGATGATTAGTGACAAAATAATAATTCGTAATCCCCCT
 AAATAAGTTGTGCAATCTGAGCAAGCGCAATACCACCATTTCCTTACT
 45 TTAAAACGACCTAACTCATCGTACCCATTAAAGCTAATAAAGTATAAATA
 ACGCCCATAGCTATAATACTAATCGTACCTGATTTTAAAGTTTCTTTAGCG
 ATTGATTCGGATTAGTAATCCCCAATTTTTAATTGTAGTAACAATGATA
 ATACCAATATGCTAATGATGCCAAAGCGCTAATGTATTATATCCATCGATA
 AACCCTTTGAGTAACACGCTATTGCTATAATCAGCACTTACTGGCGCATGA
 50 CTAATTCACCCATAGGACGGATAAATGCAAGCACAACAATTCACAGC
 AAGATGAGAAAGACCGGATTTAAGAATTTTCCAAATATAGTCTAATATTTTA
 GAAGGTTTACGCGAAAATAACCACGCTACTCCGAAGAATAAAATACTAAAA
 ATAGCAACAACGCTTGGGCCGTACCAGATGAAATAAATGGTGAAAATGCTA
 TTTCAAATGACGTCGTCGCAAGTCTTGGCAACGCAAAAAACGGACCTATAA
 55 CAAGATACAAGCCAATTGTGAACAAATAACCATATATTTTACTTATCCTTG
 AGGAAATTTCAAAGACCCCGTTGTTTTAGATACACCTATCGCTATAATTC
 CTAAAAAGGTAGTCCGATAGCCGTGATAAGAAATCCTAAATTGGCGGTCC
 ATACATTTGCCCCCGCAGTTTGACCTAAGTGAATTGGAAAAATGAGATTTT
 CAGCACCAAAAAAAGACCAACAGCATTGAGCCGATAAATAGATTTCTCTT

TTAATGTTAATTTATTTTTCATCATTAATTATGTATCTCCTAATATGTATT
 ATTAGCTATTATACATGATTTTAACCGTCTGCGTGGCACGAAGTTATATTT
 ATATATGTGTGATTATTACATCTTAATCTCATAGCTTTTCATTTTATAAC
 ACGATTCAAACATATAAATAATGAAAAATTTTAAATATTTCTATACATTGTT
 5 GTTTTTTATAAGAAATGTTAAGTTAATATTTATTTTATTAAAAATCCAT
 TTATTTTAGCATATAATTTTAAATAAATAAGTTTGTGATAATATTAAGT
 GTACAAGGATTTTATTTTATTTAAAAAGTGAAAAATCATTACTAAAAATG
 AGGAGGACTGTGAAATGACTGTATTTGAGTGAGCAAGACAAAATAAGATTAT
 TAGCTGATATCGTCAAGATTCAACTGAGAATGATCATGAAATTAAGTAT
 10 GTGAGTACTTAAAGACCTGTTAAGTCAATATGATATGATTCTAAGATTG
 TTAAAGTTAATGATTACGCGCTAATTTAGTTGCTGAAATTGGTAGTGGCG
 CTCCAGTCTTAGCAATCAGTGGACACATGGATGTTGTAGATGCTGGTGACC
 ATGATGATTGGACTTTTCCTCCATTTGAACTCACAGATAAAGATGGCAAGT
 TATTCGGCAGAGGTACCACTGACATGAAAGGTGGTCTTATGGCGATGGTCA
 15 TCGCGATGATTGAATTAACAATCAAACGCATTAAAGCAAGGTACGATAC
 GTTTATTAGTACTACAGGAGAAGAGACGGAACAATATGGTGCACAATTAC
 TAGCTGATGAAGGTACTTAGATGATGTGAGTGGTTAATTATTGGCGAAC
 CAACTAGTTAATATCGCTTACTATGCTCATAAAGGTTCTATGAGCTGTGTAG
 TCACAGCGAAAGGTAAGGCTGCACATAGTTCCATGCCACACCTAGGTACAA
 20 ATGCTGTTGATATTTTAGTTGATTTGTAAATGAAATGAAACAAGAATATA
 AAAATATTAAAGAACATGATAAAGTACACGAGTTAGACGCTGTTCCAATGA
 TTGAGAAACATCTCCACAGAAAAATGGTGAAGAAGAACACATATCTACT
 CTGGATTTGTAATGTTAACTCTGTATTCAATGGTGGTAAACAAGTTAATT
 CTGTTCCTCATAAAGCGACAGCTAAATATAATGTAAGAACTGTTCCAGAAT
 25 ATGACAGTACTTTTCGTGAAGGATTTATTTGAAAAAGTCATTTCGTCTGTGG
 GCGAAGATTATTTAACTGTAGATATACCTAGCAGTCACGATCCAGTGGCAA
 GTGATCGTTGGAGATTTAATTAAACATTTTAAAAAATAGGAGTGCATTTAC
 ATGTTTAAAAAGTTATTTGGAAAAGCTAAAGAAGTTGACAAAAACAATTAAA
 ATCTACGCACCTTTGACAGGTGAGTATGTCAAAATTTGAAGATATTCCTGAT
 30 CCAGTGTTTGCGCAAAAAATGATGGGAGAAGGCTTTGGTATTAAATCCTACA
 GAGGGTGAAGTTGTTTCCACCGATTGAAGGCAAGTCGATAATGTTTTCCTCA
 ACGAAACATGCCGTAGGGTTAAAAAGCCGAAAATGGATTAGAGTTATTAGTT
 CATATCGGATTAGATACCGTTCAATTGGATGGCAAAGGGTTTGAAGTACTT
 GTTGAAAGTGGAGATGACATTAAATAGGTGACCCACTTATACGTTTGTGAC
 35 TTAGAATATATTAATAATAATGCAAAATCTATTATTTCTCCTATTATTATT
 ACAAACTCTGATCAAACCTGAATCAATTCATATTGAAGACGTACAAGCAGTA
 GTTAAAGGGGAAACACAAGTTATTGTGATGACAGTAAGCTAATGGTTAAGA
 ATTATTCATTTTATCAATTTATTATGACCGTGCAGGAAAGAAAGACGATA
 AAGGGGTTTTTGTCTGAGCAAATTTTGAAGACCTTGCCTTTCCAAAACACG
 40 AAGATGATTTTAATACATTATCTGAATATATTGAGACACATAGCGAATTCA
 CCCTA

Sequence 3421

45 step.1002h06.cons.ok

ATACTTTTCCTAAACTTCTAAATTGATTAGCATAAAGCATCCAAATCGTCTTG
 ATGTTTATTTAGTTCAATAATCTTGTCTGCAAAATGATTAAATTTTAGGCAA
 TGAATTATTAGCTTGATATACAGCACTCTTAATTTTATTGATTGTAGGTAC
 ATTATCTTCAATTGATAGTCCGACTTTATTAGCTTCAGAAAGTAAAGCGGT
 50 TGCTACGGTTTGTAAATTGTTTATTTGCTTTATCAATCACAAACGACGA
 TCCCGTATCGGTTAACTTAGCTGCTACAGCATTAATCTTCTGATTTACTTT
 AAAATCTATATCCGCCCTTTTGAGGATGTTTCTTAAAGTACCAGTGATTTT
 ATGTGTGAATTTCTTCGGTATATAAATACCTGCATAATATTTTCCCATCTT
 AATTTTCATGGTGGCTTTTCTCTACTCACAAATGCCAGTCAAACTATC
 55 ATTCTTTTAAAGTTTGTGACCATTTTATTTCTTACATTAATATTCTTACC
 ACGAACTTTTTCACCTTGGTCTTCATTCACTACGGCAACTTTGATATGTCC
 TGTATGACCATACGGATCCCATAGCCCAGAGGTTAAACCATGCATAGAA
 TGAAGGAAGAAATGCTAAACCAGCTAAGATGACCCATACACCTGGTGT
 AGCAACTCTTTTAAATCCGTGATAAAAAAGTTTATGTCGTTTTCATTTT

CACACTCCTAATTATCATTATCTTTTTATTGTAATACATTTCCGTACAG
ATGATTACTTAAATGTGCATATTTTTTGCTTACGTAAAGATGATACACAT
ATTATATTTAAATTTGATTAAATGGATTTATTTTTCTAATGTTGATTAGTA
TAATATTTGTATTAAGGCGTAATTCATTTTTGTTGATGCTTGTCAACTTTA
5 TAAACCATAAGATAAATTATATCAAATTCATTGTTATGACAATAACTA
AAAATAATTTTTAAGGGAGCACACTGTAAAAGCCTTTAGACCTTCCTTTCT
TAATGGCGTTTCAACATATCTAAATGATAATTCGGCTAATAATACAGTAAT
TACGATATCCATAATATAAACATAAGTTGGTAATTGACCATCAATAAAATA
ACTATGAATAAAGCTAATTACAGGAAAATGCCATAAGTATAAACTGTATGA
10 ACGCTTTTCCAATGTAGACAAATAAAGGATTTCCTAATAACTTAGCTAAAAT
TGTCGTTGGATGAACAACACTTGCATAATTAGCAAAGTCATTGTTGAAAT
AAGATAAAATCCACCGTTATAAATCCAATCACTTTCATCACTAACAGTAAA
GAATAATAGAATTAAAAATGTAAGTCCCTATGATACCCGCACTATTAATCAC
AGTTTTTAAATCCCTTTAGGTGGATTGGGATTTAATTTAAAGGTGGCCAGAT
15 AAATGCTAGAAGTACACCTAAAAGCAGTGTCTGCAATCTTGTATCAGTTCC
AAAATATACTCTAGAATGGTTCAAGTGAGGTTGAGAAATAACAACCATCAT
TAATAATGAAACAGGATATGATCCAAAACATCAGTATGACATTTTTCTT
TTTTCTTAACCTATTGGCATAAATAATAAGAGTACTGCTGGGAAAAAAGGTA
AAACTGCTCTTCAATGGCTAGTGACCATAGGTGCTTTAAAGGCATAAAAGA
20 AAATTGCTCGAAATAATTGACATCTTTAGCAATATACCACCAATTAGATAC
GTAAAATATGCTGCTATCATATCATGTTTAACTCTTACAATATGCTCGGG
GTGCAATAATAAAGTTGCAATTCCAACTACTACTATTAATGCAAATACCGC
TGGTAATAACCTTTAATACGACGAATCCAAAATTTTAAAGATTTATGT
TCCAGTATCTTCATACTCTTTAAGTAATAAGCTCGTAATCAAATAACCTGA
25 AATAACAAAAAAGTATCTACGCCTAAAAAACACCTGTAAACCATTGTTT
ATTCAAGTGATAAATAATAATACCAATGACTGCAATTGCTCGCAAACCATC
TAATCCAGGCATATATCGTATTTTTTTAGCAGAGCGATATTCATTATTTAA
TTTTTTTCTCATTTGTTACATTCTTCCCTATTAATAATAAATTTCTGTTAGAC
AAGTAACTTTTTACTAATCAATTTTATAAGACGAACGCTATTGCTTCTAA
30 TCACTTATTCAAACAAATTTTATCTCATTCAAATTAACAAGGTCATTCTC
TATGTTATAGCTATCCCTATCATATAGCGCACGTTTCCATTGAGTTAATT
ATCATACTTCGTAAGATATACCTCATTAAGGTTATAAGTCTTAAGTATT
GTAATACATTTGTAATATTAATTCAATTTTATTATAGGCCTAATTTATAAA
AATATTAAATGAACATACGATAATCTTATTGAATTTCTTTAATAATTAA
35 AAAGGATAGCTACATCATTATTCAAAACAATGAGTAGCTATCCTTTATATT
ATCCGATGCCATAAGCTATTTTGGCATATCGTGACATCATATCTTTATTC
AAGGTGGACTTCAACATCATCACTTCTGTATCAGAAATTTAGGAATCT
CAGCCAAAACACTCTTAACTTGTTCAATAATTTGTGGTCCATATGGACATC
CCATCGAAGTCAATGTCAATTTCAACTGTACATAAACCTTCATCATCAACAT
40 CACTTTTATATACTAAACCTAAATTAACGATATCTATCCCTAACTCAGGAT
CTATTACCATTTCAAGAGCGCCTAAGATACTATCTTTTAGTGCTTCTTCCA
TCAGTATCACCTCTTTAAATTTCTTTACACCAATATATCAAATATCCGA
CAAAACGCCAATAAAATGCTATGATACTGTTACATAATAATTAACAACACA
AGGAGTAATATATGCAACCTTATTTAATTTGTCTAGATCTAGATGGTACAT
45 TATTAAATGACAATAAAGAAATCTCACCTTACACTAAACAAGTATTAAACCG
AATTACAACAATGTGGACACTACGTTATGATTGCTACTGGAAGACCCTATC
GCGCAAGCCAGATGTATTATCATGAACATAATATGAGCACACCTGTTGTTA
ACTTTAATGAGCATTGTGACATCATCCAAAAGCAAACGATTTTAAAGTGA
TACATGAAGTACTTGATGTGGAATTTCTAAAAATATTATTACAGCACTTC
50 AACAATCTCATATTACAAATATCATTGCTGAAGTAAAAGACTATGTCTTTA
TAAATAGT

Sequence 3422

55 step.1002h07.cons.ok

ACATATCGCACGCTGCAATAGCTACACGCTGTGCTTCTCCACCTGAAAGCAT
ATGTGGATAACAATACAAACAATCAGATTTTAACCTTACTTCACGTAGTAA
CGCTGTTACATACTCTCCATATCTCCTTTGATGTACATCGACACTGATT
CATTACTTCAAAACAAATATCTTTGACAGTATGAAAAGGATGCAAAGAAGA

AGTATAGTCTTGAAAAACCGCAGCAATTCGATGACGTCTCACTTTTAA
 ATGAACAGGCTGACCGTCCACGTCACTTGTCCTTGATCTGGTTTCTAG
 TCCTAAAAATCATACGACTCAACGTAGACTTCCACTTCCACTTCTCCAAT
 AATTGCTATTGATGCACCAGTGGGGCAATCAAATGACACGCCTTAACTAC
 5 AGGAATACGTTTATCAGAAAATATTGATTTTTTCCATAAAAATTAATAAC
 ATTTTGAACCTTAAGCATTGAAATCACCCCTCAGCACTAAGTAAAAAAT
 CATGAATTGTCTTTTAGTAGATAATAAATAAGCTGTATAGTGATGACGAG
 GTTGTGTAATACCATATTGTTTCGACCTTCTCTATCACGCGACCTTCTT
 TCATAACTACAACACGATCTGCAATTTTATTAATCACTGTGAGGTCATGAG
 10 ATATAAAAATCATGGCGCAATCAAAAATGTTGTTTATATCTTTAAAGCAT
 TGATGACTTCATATTGTGTAATGGTATCTAATGCTGTGGTAGGCTCATCAG
 CAATAATCAATTTAGGTTTCAATGCCAATGCTAGTGAATCATCATACGTT
 GAAGCATCCCACCTGACAACATGTAGGGATATGCTTTTAATAACGAGCGGG
 GATCTTTCAACTTCATATACGTCATATAATGAATGAGTTCTTGTTTTATTT
 15 GTTGTCTTGATAATGTCGTATGTGCACGCAAGTTTCAATCATTGCTTCC
 CCACTTTAGTTGAAGGATCAAATGCTCGTGTCTTGTGTCATATCATAG
 CAATATCTTTACCACGAATTTTGCCTAGTTGCTTTTCAGTTATTTGATTTA
 AATTTTGATGTTCAAAAAGATATCTCCGCTTACAGATAAACGTTGAGCAT
 TTAAACCTATTAGCTTTTACAAGTAATCGATTTTCCGCTTCCACTTTCTC
 20 CAATAATACCTAATGTTTCCACCACATTGAAGTGTGAAATTCACGTTATCAA
 CGAGTGCTCATCTGACCATGTGTCTTAATTCTTAAATGCGCTACTTTAA
 GCAATGTCATGTGCTTTTACACCTTTCTTCAATGCCAGTCGTTTCTTTA
 GAGGACATACGAGGATCAATCGCCATTTGTAAAGCATCTGATAAAAAGTTA
 AACGCCATTACAATTATGACGATAGCCACACCTGTTGTCATCATCATTTCCA
 25 GGATGTGTGAACATTACTTTTCGTGCTTCATTAAGCATCATCCCCATTG
 GCTGTAGGTGCTTAAACACCTAATCCAAGGAACGAGAATCCTGACATTTGT
 AATATCATTTGAGCACATCGAACTACTAGCAATAATCGCTATGTGAGTAAAG
 GTTAGTGGCAAAATATGTTTGCGAATGATTGTAAATCATTATACCAATT
 ACTTTGGCAAATTTTACATGATCAGCTTCAATATATTGCATTACACTGGTT
 30 CGAATCACGCGACAAAACCGCCCATCGAGTCAATATAAACGCAATAATA
 ATATTTTCTACACCCATGCCAAAACAAGTAATCAATGCCAATGTCACCACA
 TAGCTTGGAAGCTAACATCACATCGCATATACGCATAATTATGTCATCG
 ATATAACCTGGGAAATAACCTGAAATAAACCCAAAGTATCGCTCCTATCACA
 ACGGAAATAATCAATGCGACAAATACATATAACAACTAGGTCTTATGGCG
 35 TATATTATCCGTGTTAATACGTCCCGGCCATAATGGTCTGTCCCAACCAG
 TGAGACCAACTTATACCAGCAAATTTATTTGCTGTATCAATGTGATTGCT
 TCATAAAATGTAATTAAAGGAGCTAAAACCTCCAAGCACTACGTAGATTGTA
 ATAATAGCTATAGCAACTATGGCACCTCTATCTTGAAATAGACGACGTAAC
 ACAATCATTTTCAATTGCTTCCCTTAAACGCGGATTTAATATTGCATTAATGA
 40 TATCTGCAAGTGATTAAAAAGAATAAACACACCGCGACAATGAGCACGT
 ATGCTTGAATTACTGGAAATCATGTTTCAATGCTTTAACAATTAAATG
 GTCTTAAACCTGGCCACGCAAATACATTTTCAATAACGACTAAACCTCCCA
 TGATCATTTGGTATCGACATACAAAAATAGATACGGCTACCTGTAATGCAT
 45 TACGTAATACA

Sequence 3423

step.1002h08.cons.ok

CTGGATGAACCTTCAGCCATAACTCCATCAGCACCAACGGCTAAAGCCGCTT
 50 TCGCTGTCGGTAACATAATATCTTTTCTACCTGTACTATGCGTTACATCAA
 CCATAACAGGTAGATGTGTTCCCTTGCTTTAAATAGGTACAGCAGAGATAT
 CTAATGTATTTCTAGTTGCTTTTTCGTATGTACGAATACCTCTCTACATA
 AAATAATATTTTCGGTTACCTTGTGAAGCAATATATTCTGCTGCATAGATAA
 ATTTCTCTATCGTTGCTGATAAACACGTTTCAATAAAATAGGTTTATCTG
 55 TACGACCAGCTTCTTTGAGTAACTCAAAATTTTGCATATTTCTAGCACCAA
 TTTGGAACATCCAAATAATCTGAAGCGATTTGAAATCATTGGGGTTAA
 CTATTTCACTTACTACATTTAAGTTATATTTATCTTTAACATTTTCAAAA
 TTTTGAGACCTTCTACTCCAAGACCTTGGAATCATAAGGAGAAGTTCTCG
 GCTTAAACGCTCCGCCACGAATAAATTTCTCGCTTTAGCTGCAATCTT

GTGCTACAGCATCAACTTGTTC'TTGTGATTCTACAGAACAAGGACCAAATA
CAAATGATTTATTACCATCACC AATGATTCACCATTATCAAATTTAACGA
TCGTATCTTCAGGTTTCAATTTACGTGATACATAAAGATGTTTTTCATTTT
CTGATTTTTGTAAATCTGTAGAGGCTTTAAAAATTTCTTTAAATAATTGCT
5 TAATCACATTGTCGTTAAATGGACCTTTATTTTTATCTAATAGCTGATTAA
TCATTTCCCTTTTACGCTGAGGATCATATACTCGTGTACCTTGCTTAATTT
TTTCTTCACCAATTTTCTGTGCTAATTCTCCTCGCTTCGATAGTAATCTA
ATATTTGATCATTTAAGGAAACAAATTTCTTCTCTATACTGTTCTAATTTAT
TAGTCATTTCTACTCACCTCTAAGATATAGTTACGCATAAAGTATGTTACT
10 TTAAAGCGATAAAAATTTAAATTTATTCAAATCTAATTTTAAACAATTCACAT
TTATAGTTCAAGCGTCAATTTACAGAAAATACTAAATCTTCTGTAAATTTA
TGTAAGTAATAATTTGTTAAATAATGATATTACAATATTATTTAAATGTTA
AAAAAGATAGGAACGAACGAATATAAATTCGTCGTCCTACCTCAAGATTT
CAGATATTATTTAAATATCTAATTTTAAAGTGAAATAATCTCTTAAATTAA
15 TGACTTAAATTTAAATTTAACATAATAATTACAATAGTTAGTAAATCGTG
TTAAGTAGATAAAGAATATAACTTTTATAATGTTCAATACACTTCATTTTA
CTCTTTATCATTTGTTAATATTTAAAACTCGAGATTAAATCATTAAATGTTTCG
TTTTTCAATTTTGTCTACTAGTTTCTTCTACTTTTTGTGTTTATTACGTTT
GTTTTGGTTAGGTGTATGTTGCTGTTTGCCTTTGGAATCCTCTTTAGTTGC
20 TGTTATATTTCTTAGTCTGATTCTGAATTTTTTTGGCTGTGCGGATTATTTGA
ATGTTTTTTTAGATTCTAGTAATAAGTGTAGGGATACGCTTCAATTTTGTGTTG
AGCTTTAGCAATTTGATGGTTC'TTTAACAACGCTTCCTGTAAGTTGTGTTTC
TTTAGAGCCATTTGTGAGCTTAGATCTTTTTTCTTTTGTGCTGCCTGTGCTAA
GCGATAAGCAACTGAATTAGTATCTGATGCTTGAGATTGATTTAAATTAGT
25 TACAAGATTTGGAACATCAGAGTTATTTGAAGCTTTTATAGCTTCTTCTTG
CATTAGATTTTTCGTGTTACTAGCAACTTCATTATTATTATTATTCTATC
TTTTTTTGTCTTTTGCCGCTGCAGCTAATTTTTCTGTTCAGTTTCTGACTT
ACTTTTAGTATCTTCTCCACTTGCAGAAAGGTTTGTAAATCATGATTGCG
TGCTTCTTCTTTAACTGCCGTTTGAGCAATTGCAATTTCTTCAGCAGAAGG
30 TTCTTTCTCACTATCTATGTGCGTCTGATAGTGTATTTATTGTTTGAATCAAC
TTCACTTTTAGCTTCTTTGAATTGCTGCGCTTCCGTTGATTTCATCCGCTAA
GTCTGTTTCTGTTGAATAGCATTTTGTGAGCTGCTAATTTCTGTTGCTGA
TGGTCTCACATCATCCACTTGTGTTGTTTCTTTTTTTGCTTCTTGAATTGC
TTGCGCTTCCGTTGATTTCATCCGCTAAGTCTGTTTTCGCTTGAATAGCAGC
35 TTTTTGAGCTTGTAAATTCATCTGCATGGATATTATCTTTGGAATCATATTT
TAATTGCTCAACTTTAGATTTAACAGAAACCACTGTATCTTCCGCTCGTTG
TCTATAGTTTTCAACTTTACTACGTAAATCTTGTCTTTTTCTTTCAATTC
ATCCGTTTTTTTATATATTTTATTTTATAGTACAATCCAACAGCAGAACC
TACTATCGCACTTGCTATGAACTAACTACAAAATCCTTACGCTTCAATGA
40 GTTAGATTGTGAACGAGATGTGCGATGATACAATTTCTCGCTTGATGGCT
TTTCTCGAAATTATCACGATTATAATGTTCCATAAACTTTCTCTCTAAAAC
AAAAGCACTTAAATAAGTAAAGCCTCATTTAAGTGCAATTTGTCAATTAAT
AGTATTAATTTATTTAATACAGTGAATTGTAGCTATGATTTGCATCATCT
GCTACAGTGTGTTGTTGTAGTTTGCACCTTCTCTACGATTGTATCTGTTT
45 TGCCATTTGTGAGCAATTTCCATTGCAACATTTGACCAATTGTACAACCTGA
GAAATTTTATCTTCATTTTGAGAAATATTGTGCGTAATTGAGTTTGTAACCT
CTATCTACAGATCCATTTAAGTTTGTACTGAGTCTCCAATGCCCTTAACA
GCATCAACTACAGAATTTAAACGATCAACTTTACCTTGAATATCTTCAGTT
AAGCGATTCACTTTATGAAGTAAATCAGTCGTTTTCACGAGTAATACCTTGA
50 ACTTGACCTTCAACACCATCTAGTGTTTTAGCCACATAGTCTAAATCTTTT
TTAACAGAAATAAGTACCAGACGATACCGATACATAAAATTTAAACGCT
ATCGCAGCGATGATTCTCGCAATTGGTAAATCCAAATCCATTAAAAACGCC
TCCTAATTACCATAAAATAAAAGTCATTAATATTTAATACCCATGATGCTT
TTATATAAACATATTAAAAATCATTTTTTACGCCAATTTTTTCGAAATAAG
55 CTTTTAGTAACTTTTGAATGTCACCAGCGCCATAAATAAAATTACAGCGT
TATCAAATTTTCTAATACATCAATACTATTTTCATCTATAAGCGTCGAGC
CGTCAATACGATTGATTAAATCTTCTATAGTTAAATCTCCCGTATTTTCTC
TTATTGAACCGAATATTTACATAAAAAATACTTGGTCT

Sequence 3424

step.1002h09.cons.ok

5 GAGACGTCCAAGTTGAGGATATGTTTCAGTATTACCATGAGTCATACCCAT
ACCGCCACCGATAGTCACATTAAATCCAATTAACATCATCTTGTTCAACGAT
AGCGATTAAACCAATATCTTGAGAATAGACGTCAATATCATTAGATGGTGG
TACTGCAATACCTATTTTGAATTACGTGGTAAATACGTATTCCCATAAAT
AGGTTCC'TTTTCCTCACTCGAATCTAAAACCTTTTCACCATCAAGCCAAAT
TTCATGATATGCATT'TGTTCTTTGGAAGTAAGTGATTACTTATACGCGTTGC
10 ATAATCATTAAATCTCCTTATGTACTTGAGATTGATAAGGATTAGGATTGCA
CATCGTATTACGATTAACATCTCCACATGCAGCAATAGAATCAAGTACTGC
ATGATTAATATTTTTTCATTGATTGTTTCAAATTACGTTTAAGAATTCATG
AAATTGAAATGCTTGGCGTGTGTTAATTTAATCGTATGATTGTCATATG
ATTAGAGATATCATCCATAGCAATCCACTGTTTCAGGAGTCGCTTTCCCCC
15 AGGTACACGAACTCGAATCATAAACTATATGCAGGCTCAAGTTTTGTGTT
ACGACGCTCATCTCTTAAATCCCTGTCATCTTGCAATATAACTTCCGTGAAA
CTTTAGCAGTT'TTGTCATCATCTTGTTGTAATGGATCCAGTGATTGGATTAGC
TAAACTTTGTTCAATAGTTCCACGCAAAAAGTCACTATTTGCTTTTAAAAA
TTCCATTTTCATCAAGATTTTATCTAATTCTTCCGAAATATGATTATTTGT
20 ATTAACCATGTATCCCCCGTCCCTTCTATTAAATACACGTCTCTTTGATATC
TTTTATCTCTTTTCATTTGTTTAAAGTATTCTTCTGCATCTGTTTCAGATA
GGTTTTGCTCTTTGATTAACACATTTTTAATCGCTTGATGAACATCCCTTTG
CCATTTTACTTTTCATCACCACATACATAAATAGTAGCGCCATTTTCAATCC
ATCGATTAAATGTTTCATATTTCTGCAATTTTATGTTGCACATACACTT
25 TTTTATCAGTATCTCTAGAAAAAGCAACATCTAATTTTGATAAAGTCCAT
CTTCAAGCCATCTTGCCATTCCGTTTGATACAGAAAATCTGTAGTGAAGT
GTTGATCTCCAAAGAATAACCATGTATTTCCCTCAAACCTAGTTCCCTCTC
GTTCTTGCAATATAGGATCTAAACGGTGCAACACCTGTCCCAGGACCTATCA
TAATCACAGGTGTTGATTATCTTGCGGAAACTTAAATTCGGATTTTCGTT
30 TTAAATAGATAGGAATTGTATCGCCCTCTTGATTTCTCTGCAAAATGTA
CTGAACAAACACCTGAACGTTCCCGACCGTGTGCTTGATATCTAAGTCTC
CAACAGTAATGTGAACCTCATCTGGTGTGCTTTATAACTACTAGATATG
AGTACTCTCTAGGTGGTAACTTTCTTAATAATTGATGTAAATTTTCAGGTT
GTAGTTCTGTGCTGCGAAGTCATTTAATAAGTCAATCAAATCCCTTCCCT
35 CAACGTAGTTTGAATCCATCTTTATCTTGAATTTTTTCAGAAAGCTCTT
CATTATCAAAAAATATCGCAGCATTTTCTATTAAACGGTTTTGTAAATTAG
TAATTTCAAAATGCGATGTAAATGCCTCTTCAAGATTTAAAGTATCTCCAT
CCTCGTTAATTAACCTTGTGTTTCTGGACTCCAACCTAATGTGCTAATAA
GTAAGTCTACGATAGCTGGATCATTTTGAGGCAAGACAACACAAATCTC
40 CCGGTTTCATATTCCTCACCAAATATCAAGTAATAACTCAATATGTCTCG
TCTCTTTGCTGAACTCTACCATTCAAATTAATATTGGTCAATACTTCAG
CATCGTAAGGATTAGACTTACTATATTTCTTTCTTTGCTGATTTTATAG
ATTTCGTAACAACTTGTTCATTTTCAGTACCAGCCGGTGTGAATCAATCG
TATTAATAACATTTGGCCATCCACTTTTCAGCGTCTTCTTCATAATCAACAT
45 CGCAATCAGTACGATGATAAAGTCTTTCTGCCCCTAATTCAGCCAAACGGT
TATCAAAATCTTTACCAGTTTGACAGAAGAATTCATAGGTTTGATCTCCTA
ATGCTAATACTGAAAATCTCACCCCATCCAATTTTGGCGCTTTACGTCCGT
GGATGTATTCATGTAATTCACAGCGTTATCTGGTGGATCTCCTTCGCCAT
GTGTAGATGTGATAATAAATAAATCTTCTACTTTCTCAAATTCCTTAGGTT
50 TAAAGTCATCCATTGATTTGAGCGTAACGTCATTTCCAATATCAGATAAAC
GTTGTTCAAATATTTCTGCAAGTCCCTGCGCATTACCAGACTCAGATCCAT
ATAAAATTGTTATCGATCTAGCTTCCGGCTCAACTGAAGGTTCTTTTCAT
GTAGCATCGCTCTGTATTGTCATCTGAGTTATGTTGTTCTACAGAATCTG
TAGATGTATTTGATTGTTGATTGCCATCAGATAGCCACTTAACCATACTT
GTTGGTTAGATGATAAAGTGTGAAGCAATTCATTGATTGTTTAGCTTGCC
55 CTTCCGTAAGGACTATTTGGTAACTTAAGTTCAATCTTATCCACCTCT
TATACTTTGTTATTCCGATATTTGTCATTATTTATATATGCTATTTTGA
AAGTTATTAATTATATCTTCAATTCTAGTATGTACCAATTTGTTATTTTCAG
TAAATGTAATACTCAATTGATAACTAAAATCGAAAAGACAATTCATAATAT

CAAAGGGCAATACAAACTTTTATTACTCAAAATTATCGTTACTAAATAAAG
 CTGTATCAGCTTTTAAACAACTCTTTATTTCTTATCGGAATACTATGT
 TTTGAAATTAAAATTTTTTGGACTATCTCTTATGGTTTATCAGCTACATGTA
 ATCCACATTCAGTCTTACTAGAAATTGGACCAACGACCAGCAGCTGAATCAT
 5 TAGAATCAAAATACGGGTGATGTACATGGAATACAACCAATACCTGGATAAT
 TTTGATCATGTAATTCATTATATGGTAAATCCTTATCACGTATA

Sequence 3425

10 step.1002h10.cons.ok
 TAGGTAAACAATGATTGATACTTTGATTGTTCATTGCTTTATCATTTAAA
 CAAACATAGTGATTTTAAATTTTATCATATAAAGAATGAAGATAAAAAAG
 ATGTATGTGAAAAGTGTGCTCACATACATCTTTTATTAATATACATCTTA
 GTTAGCACCACCTGCTACTATAGAATTTGATGTGGATGCATCATCGTAAGA
 15 ATCAATTACTTTGAATTTTTTACCGTTAACTGTAAATGTTTGTGTTTAGT
 ATTGTGATATACCCAAAAGCCATTTATTTTCTAGCTATATCCCTATCATTAAT
 TTTTAAATCCGTCATGGTCGAACACCCTATACTCATTAATTTGATTGTTAC
 TTATGTTTGGATCGTTGATTTTAAATTTGTTATGCTAAATAATAAGTGGA
 ATGTGAAATTATGTGGAATAGATGGCTCTTTTCTCATTTCTTTAGCCATGG
 20 ATAAAGTATTTTTTCTAATGATTTTTTCGTTCATATTCGTCCAGAGTACTGA
 AAAGTGTAGTCGTTATTTCTGTATTAATATACGTATCTTTTATAGACTTTA
 ATGTTTCATATGCTTTTTTAAATCATCGTCACCTTCTTTAATTAAGGTT
 CAAAATGTTTTCTATAATCTCGTAATTTTCGTGGAAGTAGTAGTAATGTAAA
 AGTAATATTTTTGATAACCTATATTTTGAGTTTTAGTTATTGCTTCATCGG
 25 TGAATCCAGTATAGTGATTTTATCTTCTTTATCTTTGAGGTATTTTTCTA
 AGTTATCCATTTTTTCCCTTATTTGCTCTATACTCAAAACCGCTTAACACAG
 TTCCCTACCATCATACTCATAGTATCTCCGTTATCATTGCTTCGCATGGAGC
 CTTCTTCATGGATTGATTTCCATATCAAAAGGGATGCTTGCATTAACACAA
 TGTCATGATCATCGCAATGAACGTAAACTTCGGCGCCATCACCACCTCCCTA
 30 CAACATTTGTGGCTTTAACATTAAGGCTAAATTTATCTTTAAAAAATGTGTT
 CGCCAAGTTTTTCGTATTTCTTCACGATATTTTTTCGCATGTGCGATAGCAC
 TTTTTTCAGCTAATGGTTGGAAATCTTGACCTTTGTATTGACTTGCAGGAA
 CTTTTTCTATTTTCGTCTTTTCATTGTCGAACAACCTCCCTAATATTAGTGTG
 AAACGATAAATTGTGGCGATTAAATTTTATCATGATACACCTTTCTTTACTAA
 35 TATTACATCTTAATTAATATGAAATACATTATTTAAATCAAGCCCAAAAGC
 TATTTAATTCAGCAATATCTCTATCATTGATTTTACTCATGTTTATGCACT
 CCATATTTCTAATGGGTTGTTAGCATCATAAAATGCGCGTGTAGAATTAATT
 TTTGGCTCACTTAATTGTAAAGTTATGTTTGTGTTTTTCAGGTTTTCTTTG
 TAATGATATAAATCATCCTTAAGAATAATAAATCTTCTTTAGAGTTATTT
 40 TTAGTGAAATCAGATTTAGTACTGAATAATGTTGAAACAACCTCAATTTTG
 GCTTTGAAATTTGTTGCTTCATTTGCTTTTTTTCATGCCTTCTTTGAAATTC
 TTAGAATCTTTTTTTATTAAGGTTTCAATATTTTCTATAATTTATTAATA
 TTAGTAACGTTACCTGCTAAATAATAATACTCATTTTGATATCCAACATTC
 TGTGCTTTATGTATTGCTTCTTTAGTAAACCGGTATATTGATATTTTTCT
 45 TCATTGCTCTTAAAGAACCTTTGTTAAGTTATCATATTTTTCTTTTGTGCT
 TTATATTTCAAAACCACTTAACACCGTTCCAACCATATTACTCATGGTATCT
 CCGTTATCATTGCTTCGCATGGAGCCTTCTTCATGTATTGCTTCTTATCA
 AAAGGGATGCTTGCATTAACACAATGTATGATCATCGCAATGAACGTAG
 ACTTCGACGCCATCACCACCTCCCTACAACATTTGTGGCTTTAACATTAAGG
 50 CTAAAAATTATCTTTAAAAAATGTTTCGCCAAGTTTTTCGTATTTCTACGA
 TATTTTTTTCGCATGTGCGATAGCACTTTTTTTCAGCTACTGGTTGGAAACCT
 TGACCTTTGTATTGACTTGCAGGAACCTTCTTCAGGAATTGATTGTGATGAA
 TGTGTTGCTTCTTTATGCACTGTGGAACATCCTCCAATTAATAATGTTGAA
 ACGATAATGGTAGATATCAGTTTCTCATGATACACCTCAATATAAGATTA
 55 GTTACTTTAATTTTAACTTAATTTATAAATTATATAAAAAGAATTTATAAAA
 AGTATTGAACAAATATAACATTTTTCACATATTAGTTATATTGATTTTATA
 GCAATTTCAACATATTGCAATAAAAAATAAAAAAGCTTACACAATGTAAT
 TTGATACATGCATAAGCTTTGTCATATGATTAAAGATTTTTATTCTGTTTA
 TCATTTGTTCATGATTTTGTGGAATTATTTTGTAGTTGCCATTT

TGTTGTGAATCAGATTGATTATTATGTTGTGACGACTGTTGTGGTGTGAT
TGTGCTGGTTGATGATATGATTGATTAGTTTTGTTCTGATTCCGTTTCTGT
TTGTTATTTTCTGAAGAGTTATTTTATTTGAACCAGACTGTTGATTATTA
TTTGGGTATGATGGTAGTTAGGTTGTTGTGCTGGTTGTTTCAGTACCTGAT
5 TGATTATCGTTGTTATAGTCATTATTTGTTGGTTGAGTTTGGTCTTGGTGT
TGATTAGTATCACGATTTTGAATTGAAGTATCATTGTTTGTGTTTTCGGA
TTGTGTTCTATTTTTTTTATCATCTCTTATTAGAATTAGTATGGTGGTTAAT
AGCTGCATATAATGAAAAGCCTAATGCACATATAAGAATGATTAC

10

Sequence 3426

step.1002h11.cons.ok

15

TAACCCAGCCACTCGAACCATAAAGAATTCATCTAAATTTGAACGAAGAT
AGAAATAAAGTTAAGTTTTTCAAGCAACGGATTATTTTTATCATATGATTC
TTGTAATACTCTGTAGTTAAATCTAGCCAACCTAACTCCCTATTGTTGTA
ATACTGCGGTAAATTAATATCTTTTTCTCCCAATCGAGTTGCATACTATA
CATACACCTCATTAATTAATCTTTATTATGCCCAATCTAAAATATCATAAT
AATTTTAAGATTTTGTAAATATAATAGAAACCTTACCTTTTAAAAATTTCT
CAATATGCTTTTTCTGACGATTGCTTGGTATCTTCTGCAATAGGTGAAC
20 CTTTGTAAATAAACTAATAAATCGTATTTGTCATCTTCTTTGCTTTTAGTT
TAACCTCCCTACAAAACCTAGTATGTGAGATATTCAAGGTGTTTGCAAAAT
TAATAATCCCTCCTAAAGCTTGATTGTATCTATTTCTTTATTACTAAACC
ACTGTGTTTCTTTGCAATAAAATTTAAGTAAAGATTGTTTTTAAACTAG
CTAACAAAGCTAATTTTACACGATCTTTATGTGAAAAGCCGTTAATCATTG
25 AATTTGCGATTAAGTAATACGTATGTGGTGAACCTGAGTCTGAATCAATGA
AACTACCTAGATAATAAATGTAGGCACCTTCAATAAATAATCTTTTCCA
TTTCTGAAATATTAAGTGATCGTTCACTTATAATTTGATTCAATAAGGATT
GAGCTAATTTACACGACGATTAGCACTCGTTTCTTCAATATGATATTCGT
TCGCTAAATGACGTAATGCATCTTTACGTACGTTACTTTTATTAACTCAT
30 CAGGATATCGTTTGTGCTGATGTGGTTCATAATAAATCCTTCACGAAATTCCTT
TTCTTGAGAAGGTGAATTTGTGTGCGTCAATTTTTTTGAAAAGTGTTTTAA
AGACGGAGATAGCCGGCAGAATAATATCGACGCGATCGCGACTTAAACCGT
CTAAATTTGTAGTTTCATCGCGAGAATTTTACGGATTAGATCATAAACAT
TGTTAATATCTTTTCAAGTCATCGTATAGTTATGAACGCCACCGATAGGGT
35 ATGCATGTGCTGATTGATGAATGCGTGCAACATTACGTGCAGAACCACCTA
CTCCTACGAGCGCAATATGTTGGTTGGATAGCCAGTCTAACTGACTAAAT
GCTCAGGTAAACCTGTTCCATATTTTAAATGGCTGTTTTGTCAATTATGTT
CTTTATCACCAAAAAAAGTACGCTTAAAGTATACCACGCCAAATGGAAGC
TATGAGCCTCTTTAAGTTGTTTGTCTTTGAAAAGGGTAACTTCGGTAGAAC
40 CGCCACCGATATCGACAGAAATCCATTTTCAATATCAGTAGTATGTGTAA
TCGCATAGTAACCGTAAATGCTTCATCTTCTCAGGTACAATTTGAATTT
CGATATGAATATCTTGTTTAATTTCTTTAATGATAGCTTCACGATTTTAG
ATTGACGGATAGCAGCTGTTGCGATGGGATATAATGCATCAACGTTAAAT
TATCCGCAACTTTTCTAAAACCTGCTTAAAGTCTCTTTTAAACATGAATAC
45 CTTTCATCATCTTTCATTTGACTTAGTGAGATATTGACTTAAACGTGCAG
GTGTTTTTATATTCAGTATTTTCATTGAGCCCAGTTTTTTTATTGTAGCCAA
ATATAACAAGTCGAATCGTGTGGAACCAATGCTATCAAACCAATTCCTT
CTTCCATGTTTGCCTCCTATTATTCAAACAATGCGTTTAAATAATCCCTA
TTTAATTTTATGTACATGATATAGATAAATGAAAAGAACTTACAGTTAA
50 TTTTCTTGAAATATTTAAGATGATTGATAAGTAATTATAATTTTTTAAAT
ACCAATTTTAGTAAAAAGAGAAGTATATAAATGTATAGCTTAACTCAGCTAT
TCATTCCTCGTGCCTTATTCTTATCACTGTTTAAAGTTGATTACATCGT
GGAAAACATAAACTAACATGATACCTAATTGTATTAAAGTATGTACAAGTGA
TAGGAGAGTGACTTTGGATGGAACGTTTAGAAAACAAAATCGCAGTGATTA
55 CTGGTGCGAGTACTGGTATTGGACAAGCATCGGCCGTGGCGTTAGCAAAAG
AAGGAGCACATGTGTTAGCGCTTGATATATCAGATCAATTAGAAGAACTG
TGCAGTCTATTAAATGATAATGGTGGGAAAGCAATGCATATCGCGTAGACA
TTTCAGATGATAAACAAGTCAAACAATTCTCAGAAAAAATAGCACAAAGAT
TTGGACATGTAGATGTTATTTTAAACAATGCGGGTGTAGATAATGGCGCCG

GACGTATTCATGAATATCCAGTTGAAGTGTGGATAAAATTATGGCTGTG
 ATATGAGAGGAACCTTTTTTAGTAACTAAATTTTTATTACCTTTAATGATGA
 AACAAAGGTGGTCTATTATTAAATACAGCTTCATTCTCTGGGCAAGCTGCGG
 ATTTATACCGTTTCAGGGTATAATGCTGCTAAGGGCGGTGTCATTAATTTTA
 5 CAAAATCTATCGCTATAGAATATGGACGTGAAAATATTCGTGCTAATGCTA
 TAGCACCTGGAACAATCGAAACACCACTTGTGTGATAATTTAGCAGGTACAT
 CAGATGAAGAAGCCGGACAAACATTCGAGAAAAATCAAAAATGGGTAACAC
 CATTAGGTCGACTAGGAACACCGGATGAAGTTGGGAAACTTGTAGCCTTTT
 TAGCTTCCGATGATAGTTCATTTATAACTGGTGAACATTTTCGTATTGATG
 10 GTGGCGTGTAGGCTTATACATGGCTACAACACGCATTTTCATTTTTTGTGCG
 TTGTTTTTTCTATTCTTTCTGTGACGTATATTGGTTACCGATATGTTCTAC
 TTTTTTATTCATTGCTTGTACCTCCTTAAGCATTTCACTCTTCATTAATA
 CGTTCTTCTTAAATC

15

Sequence 3427

step.1002h12.cons.ok

TGTTCTAAATTAGCAATAACACTTTCAAGAGACGCACGTGTAGGATTGCA
 GTACGTGAATATTCGTACCCTTGCTTAAATCACCAATATCATCTTGTA
 20 TAAGTACTTGTGATAAATAGGTGTTGTCACGTCTCCAGTATAGTTGTCT
 GTCGTATGTCCCATGTATCATTTGCGTTTTTTTATTTCATTATTAAACT
 CTCCTTATAGTTGAATATTTGTTGGACATGTATCGATCGCTTCCATCTGG
 AAAGATGGTAACAATCACACCATTTTGAATGCTTTTTTTTAATTCCAACGC
 TCCTTGTAATGCCGCTCCCGAAGAACTACCAACTAACAATCCTTCTTTATT
 25 CGCGACAAGTTTAACATTATTAAGCATCTTTATCGGCAACAGTAAAAAT
 ACCATCTACTAATTCTTTTTCTAAAAATGATGGCCACTTTTCAGAACCAAT
 CCCTTCAGTTGCATGAGGATGACTGACACCACCATTTAAGACAGAGCCTTC
 TGGTTCTACAATATAATTTTTTACATCATACGTTTTTAAAGTGTGTGCAAC
 TCCTGTAAACGTACCACGGACCCAAACACCTGCCACAAAATAATCAATATG
 30 TGAAAGTTCATCTGTGAGTTGTTTGGCAAGTGTGTGTATATGCCCGAGG
 ATTATCTTTAGTTTCGAATTGATTCATATATAAATATCCATATCGTGTGCG
 GTATGCCAACGCCTCTTGCTGTGCGCCAGTCATTCCCTCAGCTTTGGGGGT
 ACGTCTAACATCTGCACCCAAATGCTTTCATAATTGAAATTTTTCTTCTGC
 AAATCCTTCTGGAGCAAAGATGATACATTTTACTTTGTGCCGATTAGAAGC
 35 AATAGCAAGTCCAATGCCTGTATTACCAGCAGTCGCTTCGACTATTGTATC
 CCCTTCTTTAAGTCGTCTTCATCTATTGCTTTTTCAATTAAAGTACTTCCC
 TAGACGCTCTTTGATGCTACCACCAGGATTAATTTGCTCAAGTTTGGCGTA
 TATTTTAACATTCTCGTCACTAAAGCTTTCTAATAAAAACTAATGGAGTTG
 TCCTATCAAAATCGTATGCAATCATTATTTTTGGTTCAGATGACAACAATGA
 40 GTATCACTGAACTTAACCACTTCTTTATTTATGTTAGCGCTAATCTTA
 CTTATCAGATATGAATTATAACTTATCGGCATTGTTTTGCAATTTCTGAA
 AAACGTACAATCAATTTGAAATTTAAACATTATTTAAAAATAAGTATTTTA
 TAATTGATTACGTTTATATGAGAATACATACCGAAAGTGTTATGATAGGGG
 TAAATAAGCTAGCACATCGATAAAAAGTTATCGAATTACATACACTAGAAAT
 45 GATAAGGAGCTGCTTTAATTATGACGAAACAAGATTTATCTTTATCTGTAT
 TTACCAATGAAAATTATAAGAATCTTCGTTACACATCATCTAGTTTTAGAA
 ATTCTATGTATGATGAATTAGAGGTTAATAAAAGTCGTTTTAAAAACTGTA
 ATTTTAATGAAGGTATTTTAAAGAAATATAGAAGCAATTTGTAATTGTAAAT
 TTACAACGTGCGGGTTAATAATTGTATTTTCGAAGATGTTTCATTTTACA
 50 AAAACCAATTTAAAGATTCAACATTTGTGAATACACCATTTGATCAATCCG
 TATTTAATGATGACTTTTATTTCCAAAATGCAATGTTTCGATAGCAATCTCATTC
 GTAGCGTAAAATGGACTGATATCATTTTAAAAACGTTTCTTTCAAAAATG
 TAGAAATGAAGGAACAACATTTAAAGATGTAAATTCAAAAATTTGTGAGT
 TCAAAAATGTAATTATTACTAATTCAACTATGTCGCAAAAGTTAATGAATG
 55 AATTACAAAAACAAGATGTTACTTTAGAAAATATAGACACTTCTATTAAAC
 ATTCATTATCTTCCGCTTAAGAAAAACAATTTAAGATCATAGTTAAACATTC
 GCTCGTTTAATGGTGTTAGATATTTCAATCACTACATAACATTTCTTAT
 ATAGTAAACGGTCAACTCCACTAAAAAATGGAATTGACCGTTTGTTTTAG
 TTTTATCAATAACGTCTCAATGACTTAAAAATAATTGAACATATAAATACTAA

AATAATAACAATAGGCATGATGAACTTAATTAAATAATACCACGGTAAAAA
 TAGGTTGAATTTGTCTTTCCCAAAGCTTTCTTTTAATAATTTTTATCTAG
 TAATTGGCCAACCACTAATGTTGTTCTTAGTGCCCCAATGGCATAAGAAT
 ATTAGATACAATAAAATCCATATTATCAAATATCGTACCAGCGCCAAAACG
 5 CAAATGACTTAGACTACTAAAAGATAATGTTGCTGGGATACTAATGATAAA
 TACAAGTATACTACCTATGATTGCCACTTTTTGTCTTTTACTATTATCATT
 TTTAGTAAAATTAGATACATTTAACTCTAATAATGATATAGAAGACGTTAA
 TGCCGCAATAAGAATAATAGTAAAAATATAAAGTAAAAGAATGTACCAA
 AGTCATTTCTGCTAAATACCTAGTGGTAAAACCTTAAATAATAAGCCAGGGCC
 10 TTCTTGGGGTTGGTAACCAAATGTTTAAAGCGCAGGAAATATAGCTAATCC
 AGCCAAGACAGAAATTTAAATATTCATTACGACAATTGAAAGTGCTGAAGA
 CTTTATCGTCAATTTTATAGGTGCATAGCTTGCATAAGTAATCATTCGGGT
 TGTACCTAGGGACAGCGTAAAAAACGATTGTCTTAACGCAATAGTACACC
 TTGAATAGACATATCTTCAACTCGAGGTTGCAGTATATAACGTACACCTTC
 15 TAAAGCACCTTCTAAAGTTAAAGATTGTGCTACAACGATAATTAAAAAGAT
 AAATAATAGAGGCATCATTATTTAGAAGCTTTTCTAAACCTTTTCAAC
 ACCTAACATAACAATTACCATTGTTATCAAAATAAATATGCCTTGACCTAA
 AACAGTCAACCATGGATTACTAATGATTGTTTCAAATGAATATTTGTCAG
 CGTACTTGATTAAAAACCATGATTTGTGCTATGACATAACCTATATA
 20

Sequence 3428

step.1003a02.cons.ok

AAAATCAAATCCGGTTTAGCTTTTATTAATTCCTTTTATTAAATTCATC
 25 GCATCGAATTGTTTTTTTACCTTTTATTACATCTTTAGGATAATCATCCACA
 GTGGATACACCAACTATATCTTCTCCGATTCTTAAGCGATATAAAATTTCT
 GTGTTACTAGGAATGAGCGAAATAATTCTATGATATTTTTATCGCTTTT
 TCAGTGTCTTTATTATTAGATGAATGATTGGCATCTTGACCACATGCAGCC
 AGTACTACAATCAATGCACATATAACTATCCATAGTTTACTTTTCATTG
 30 AGCATCACCTTTCACAAATATTTTTAATGTTCCATTCATCATATCAAA
 ATATACCAATATTTGTAGATACAATTTTATTGATTTCACAAATATGAA
 AAAAGACAACTTTGAATTAATCTATTAATATATTTAACCCCAACTTGCTT
 TGTCTGTAGAATTTCTCGTTGAAATTTCTGTGTTGGGGCCACAACCCCAA
 CTTGCTTTGTCCGTAGAATTTCTCATTGAAATTTCTGTGTTGGGGCCACA
 35 ATCCCAACTTGCTTTGTCCGTAGAATTTCTCGTTGAAATTTTCTGTGTTGG
 GGCCCCGACCCCAACTTGCTTGTCTGTGGAATTTCTCGTTGAAATTTTCT
 GTGTTGATCAGTTTATAATGTTAGTATCAATCGGTAAAATGCTAAAATGT
 CTGAGACAATCAACAATGTCCAGACACATAATGTTTATAAATTTAATAAA
 TGTTTTTATCGCGTCAGAAATGTATAAATTCAAACCACGTTACGTTGCTT
 40 GTTAATAAACTTTACAAATCGCTCGTCTCAATATAAATCTGAGCTATATA
 TTCTAAAAATGATTGTCACAATTAGGAACTTGTCTTTCAAAATACATTG
 TAATTTCTTAGTTTTTCCACTTGCTTCAAGTATAGAACTTTGTTTAAGTA
 TAGTTGATTCAATTCGTCAAAAAACATATTTAATTGCTTATTAATTTCCCTC
 AAAATGATTTTGCTGTTCTGATTCATCCTTACATTTTGTGTTATCTTTATA
 45 TGCTTGATAATAATGAGTATCTCCATACTTACTTGCTGCTTCTTTTTCATA
 CTGTTCAATTAATAATGAAATTTTGCAATTGTAGTTACCTCAATCTCATCGTT
 ATCTTTGTGAAAATGATGTTCTAACGTACGGATAATATCGTTTAAGCGATC
 TCTTTTTTTAACAATGTGATAAAAATTATCATGTAATATTTTAGTTCTAAA
 ATCACTATCACCTTCAAAATACGTCTGTATTTGCTTTAATGTTAAATCTAA
 50 TTCTTTTAAAAAGAGAATATATTGTAACCTTTGAAATATCTTGAACATTATA
 ATATCGGTATCCATTAGCACTCATTTGTTGTGGCAGAGTAACCCCTTTTC
 ATGATAATAATGAAGCGTGCGTATACTAACTCCGTAATGTGAGAAAGCGT
 TTTGGGACTGAAGCGTTCTTTTCATCATTTTACCTCCTGAATTAAGTATGA
 GGTATGACGTTACGTTAAGGTCAATACTTATTTTTTAAAAATAGGTATCTAA
 55 AAAGGTATAGCGTTGTTTCCCTTACCATTAAAAATAATTTTCCCAAAT
 ATCTAATAAAGCGTGAAACTCATTTGCATCTTGATTGAAAATGATTGAGG
 AAGTGTAAATTCCTGTTTCAATTTATGATAACTTTCTGTATGTTGATATCC
 CAATTTTCTAAAAATACGTCTAGTATAACTATCAGGTATGAATTTCTTACC
 TTTAAAAATATATACTATTAAAGACATCGGCGGTCTTTCACCTATACCACG

GATGGTGAGTAATTCCTTTCTTAAGCTATCACCGTATAACTTAGCTATACT
ACTATAATCAAAATGATGTTGATTTAACCATAGTAACAATGCCTGTATAGC
CTTAGCTTTATTTTTATAGAAACCACTCGAACGTATCACTTGCTGCAACGA
TTCTAAAGGCATTTTCAATATCGTCTGTGCATTAAAAGAAGTTTCTTCTTT
5 TAATCTTGATAACGCTATATCTGCATTGTTCCAATTAGTATTTTGGACTAG
AATTGCCCCAAGCATCATTCTATTGGCGTTTCTGCTGGCCACCAATATTG
AGGGCCCATATTGTATTTTGAGGAACAATAGGTATCTATAATGATTGATTT
AACAATGTTTGTAAACATGTTAACAGATTTTACAAAGACGGTAGCTAAATAT
ATAATATGGTACATAGATGATGAAAGTGAGATGACATTATGGGACGTAAAT
10 GGAACAACATTAAAGAGAAAAAGCCCCAAAAGATAAAAAATACTAGTAGAA
TATATGCCAAATTTGGTAAAGAAATATATGTAGCTGCAAAGTCTGGTGAGC
CTAATCCAGAGTCAAATCAAACTTTAAGATTAGTATTAGAACGTGCAAAAA
CATATTCAGTACCTAATCATATTATAGATAGAGCTATTGATAAGGCTAAAG
GCGCTGGTGACGAAAACCTACGATCACTTAAGATATGAAGGTTTGGTCCGA
15 ATGGTTCAATGCTTATAGTTGACGCATTAACAAACAATGTAAATCGTACAG
CATCAGATGTACGTGCTGCGTTTCGGTAAGAATGGAGGAAATATGGGAGTAT
CTGGTTCACTAGCTTATATGTTTGACCATACTGCAACCTTTGGTGTAGAAG
GTAAATCTGTAGATGAAGTCTTAGAAACACTAATGGAGCAAGATATTGATG
TAAGAGATGTAATTGATGACAATGGCTTGACTATTGTTTACGCAGAACCAG
20 ATCAATTTGCACAAGTTCAGGATGCATTACGTGAAGCTGGCGTTGAGGAAT
TTAAAGTAGCAGAGTTTGAATGTACCTCAAACGTATATTGAGTTGTCTG
AAGAGGATCAAGCTATTTTGA AAAAATTAATCGATGCACCTGAAGACTTGG
AAGATGTTCAAAAATGTTTTCCATAATGTAGATTTAAAATAATGATGACTGC
TAATGATTGGATAGACCGGTTGAATTAATTTTCGCATCCTGAAAGTTTCAA
AACGATTTTCGATCAAACCTGTTTATTTGTGTCAAGAAATGTTCCATCCATAT
25 CTACCGCAATGGCTTTAATCATCGTTAATCCCCCTATCTATAGATTAAAAT
TCATAATCATTATATAGCTGTTGGATTTGGTATTGAAGCAATATCAAATAT
AGTGA AAACATATTTCATATATTGAAATTTGTTTCAACATAAAAAATAATTA
CCTTTGTCAATTCACACTATCGCTTAATATATAAGATAAAATCTCTTATTTTC
30 ATTTATATAAATTTGAATCAAAAAATTTAAAAGAGCCAAGGACATTCACCT
TGTCCCTCGACCTCTAATCAATAAGTTAATTA AAAATTTATTTTAAACAAAGCA
TCTTCTCGAAGAAGAACATAATTAATACAGTTTATTACGTATCTTTCACTT
ATAAGAATATTAGTTAATTTGTGTAGCTATGAATATTTTCTTGTTCCTCA
AACACCGGTTCTTCGTTATAGTCATCGCAAGTAATTAATGTTAATTGATT
35 TTTCTTATTAGGATGTTTCGTCTAATACCTTAACCTCTGTAGGCTTAACATC
ACGTATTTTAGTTATTTTATACTTTCTAGTTTGATTTCCAGTTTAAAATA
CACTTTACTACCGATTTTGGCTGATTTTAAATTTGTAAATTGATAGTGCGA
ACGATCTGTAAACGTATGACCAGCAATTGAAATATTCTGTTGATTAAAGAGA
TTCGTCACCTTCTGCAAACTAACACCTCTATTGAGTTGTTCTGGTGTGTC
40 TGGACCAGGGTATACTGGTTCTTTTATTTGTGCATCTGGAACCTCTATATA
ACCAGCCATTTTAGATTTATCGGAAGGTATCTTTGGCGTCGATTTAGATGT
CTTTGTCTGTTCTTTTTCCTTTTATCATAATTTTCAATTTTATGATCGTT
ATCTTTTTCATGTAGATAATTATCGATATATGGCTTTGA

45

Sequence 3429

step.1003a04.cons.ok

ATAGTGTTGATCTAGTGCACCTGCGGATCCAGTATTAATGACGACTTCTG
GATTAAATTTTCTATTAAACAAAGTCGTTGAGATAGAAGCATTAACCTTTAC
50 CTATACCACTTTGTGTTAAACCAACCTCTTTTGGGTTTAGCTTGCCAACAT
AAAATTTAACATGCGCAATATTTATTTTCATTCATATCATTCAATTTACGCT
TTAAATCGTCACTTCTTCTTCCATTGCTCCAATAATTCCTATCATATATC
TTTTTACACCTCATTTATAAAGCATTAGCATAAGTATTTTATCACATTTTA
AATATCTCCACGATAGCTAGTTTAAAGTATCAATGCTTAAATCACTATGA
55 TTAAATAATATCTCTTCTTATTTATAATTA AAAACTACCTTGCAATTGATT
TAAAGACAGTTGCGGGTAGTCATAGAACCTTTATFATACAATTAACCTTTCC
AAAGACTTGCTAAATCTTGTAATGAAAAGATACCAGTCATTATTGTGACAA
TTACGGCTATAATACCAAAAAACAAGCATCCATGTTGGGTGTTGGTAATTAC
CAACGATAGATTTTTTCTACTTGCAATGAGAATTGCACCTAATGTGATTG

GTAGAATCCAACCATTAATCGCACCAGCTATTATAAGTAAACTCACCGGTT
 TACCAATAAATAAGAAAAACAAAAGTTGAAATTACAATAAATGTAATAACGA
 TAAGATTATTTTTATTGAGTAACGATTTGTGTAGTGTTTTTAAAAATGTTG
 CGCTGTATATGCAGAACCAATTACTGAGGACATTGCTGCTGCAAATATTA
 5 CTACGCCAAAAATATTTTTACCTATAGGACCTAATGCATGTTGAAAACTG
 ATGCTGGTGGATTTTCTGAACTAAGCGTAACGCCAGTTACAACACACCTA
 GTACAGCTAAAAACAATAAGGTGCGCATGACACCAGTTGTTAAAAATACCTG
 CTACAGCAGATCGATTTACGAAAGGAAGGTATGACTTACCTTTTATACCAG
 AATCTAGAATTCTATGTGCACCTGCAAAAGTAATATAACCCCTACTGTAC
 10 CACCAACTAATGTAATTATAGGTAATATAAGTTTGAAAGGATGTTTCAGGTG
 CAAATGTATGTACTAATGCATCTCCATAAGGGGGATTGAAACAACCATGA
 CATAAGCGACTACTAAAATCATTACGATACCTAGAATCATACTAATAACAT
 CCATTATTTTCTGACCATTCTACTAACAAGATAAGTATCGCAAAAATAG
 CTGTTATTGTCAGCACCCCATTTTACATCAAGACCAAAACATTGCATTTAAAC
 15 CTAAACCTGCACCAGCAATATTACCTATGTTAAATGCGAGACCACCAAATG
 CAATTAGTATGGAGATAATAGTACCAAGTCCAGGTAATACTTTATAGATA
 TTTCTTGTCCACGTAATCCAGTTACCACTAATATTCTCCAATATTTATTT
 GAGCGCTATATCTATAATAATAGAAATTAATATTGCAATGCAAACTAG
 CATAAAATTTGTGCAGTAAACACTGCAGTTTGAGTCAGAAATGCTGGACCAA
 20 TAGCTGATGTAGCCATTAAAAAACTGAACCATATAATAATCGTCTATGTT
 TTTTGTGAATTTAAAAATTTGTCTTTTGTGATTGAAATCTGTTTGTAT
 TTTCCCCCATCTTTTGACCCCTATAAGGATTGAATATCAATGCCTTCTTT
 CATTAATTCATTTCTAATTGCGTAACAAATCTAATGCGTGTTTTCCATC
 TCCGTGAACACAAATTTGTATCAGCTTTCAATCGATGATTTTCCATTTTT
 25 ACTTACAACCTTTATTTTCCAAAACCATTTTTAATGCTTGTGGATTGCTTC
 GTCAGTATTAGTGATAGTGGCATCGGTTTTTTTTCTACTTACCAATGTCC
 GTCATCTTCATAACGACGGTCAGCAATACCTCCGAAGCTACCTTAAGCCC
 CACTAATTCAGCTTCCGAAATAAGTAATGTATTCGCTAAGCCGACGAAAAT
 TAGATTTGAGTCGAAATCAAAAACCTGCTTGAGCAATTGCATGTGCAATTTT
 30 TTTATTTCTAGCCCCCATTTGATAAAGGGCACCATGAGGTTTAAATGCAT
 CATTTTACATGATTGATTTTACAAAAGCCACTTAACGCACCTAATTGATA
 TATAACTAAGTTATAAATTTCAATTTGGAGTAAGATCCATTTTACGACGACC
 AAAGCCTTTCAAATCAGGAAGACCCGGATGTGCTCCAATACCTATATTATT
 TTTTTTCGCTAATTGAACAGTTTCATTCATAACATCTTCATCACCTGCATG
 35 ATAACCACAAGCAATATTAGCAGAAGTTATTAATGGTAAAATGTGCTGATC
 TCCACCGAATGTATAGTTACCAAAAGCTTCTCCTAAATCACAATTTAAATC
 AATTTTCAACTTCATTTTCACTCTTTAATGATTTGATGTTTTTCAAGAAAT
 TTTATATCTACTTTACTGGCATCGCCTTTTGCATATATTGGATAATTTAAA
 ACCGCATATAAAAAGTCAGCTGTAGTAGTAAACCTTCGATGACCATTTCAT
 40 TCTAACGCAACTTTTAAATTTATTAATAGCAGTTGTCTATTAGAATCCTTT
 ACAATCACTTTAGCTACAAGTGAATCATAATAAGGAGAAACCTGATAACCT
 GTGTAAAGTAAAGAATCTACACGTATATTAAATCCTTGTGGTAAAGTGAAT
 TTATTAACCTTTTCCGTGGAGAGAATTAAATGCTAAACAGTTCAAGCTGGTG
 CTTACGATATTCAATTCGTAGACAATGGATACCAATACAACCTCACTTCAA
 45 ATGGTTCTGAATGGTCAATGGAGCTACGCTGTAGCTGGTTTCAAGTCTGATT
 ACACAGAATCATCATCAAACCAAGAAGTAAGTGCAATACACAATCTAGTA
 ACACAAATGTACAAGCTGTTTCAGCTCCAACCTTCTTCAGAAAGTCGTAGCT
 ACAGCATCAACTACTTCATACTCAGCACCAAGCCATAACTACAGCTCTC
 ACAGTAGTTTCAGTAAGATTATCAAATGGTAATACTGCTGGTTCTGTAGGTT
 50 CATATGCTGCTGCTCAAATGGCTGCACGTACTGGTGATCTGCTTCAACAT
 GGGAACACATCATTTGCTAGAGAATCAAATGGTCAATTACATGCACGTAATG
 CTTCAAGGTGCTGCTGGATTATTCCAAACTATGCCAGGTGGGGTTCAACTG
 GTTCAGTAAATGATCAAATCAATGCCGCTTATAAAGCATATAAAGCACAAG
 GTTTATCTGCTTGGGGTATGTAATCCAAATCTAATAATATAATTAGAAAGA
 55 GTGAGACAGCATTCGTTGTCTCTCAGCAACGAGGACTAAAAATTGACTCCT
 TGTATATAAGTAAGTTTAAATTTTAGTTATCGATGCAATACATTAAAAAGC
 CTAGGACACATTGCTTTGTGTTCTAGGCTCTTTTTTATTTGAAAGTCTAAT
 TAAATGATTTTATATAACAAAGTGATGTTAACAAAGTAAAAATATAACATC
 TAGACTGAAAAAATATAATAAAATTAATGATATTATGTTGTATTATTT

AAAAGATATGGTTATATAAAATGGTGTGTTAATATCATATATAGATACATT
TAAATTGCTTGCCTACACAACCTATATGAAGATGCTCTAATAAATGTATAC
CTATTTAATTTGTTAAAGTATAATATTAAAGTAAAAAGATGACATGTCAAT
GCAATTGAGGATGTAATTGTGATAATATTAAGA

5

Sequence 3430

step.1003a06.cons.ok

GTATTATATTCAATATTGACGACATAGTCATCTTGATCTTCTTCAATAATA
AATGACAAATCAAATTTAGCTGTTGTTGACTGAGGTGGAATATGTGTCAAT
10 TGACTATGTCCAAAATTCGCATGATTTGTTTCATTATTTTGAAGTACGAGC
ATCACATCAAATAACGGATTATGTGAAGCATCTCTTTCATCAACAAGATCA
TTGACTAAGCTTTCAAAGGATATTCTTGATGTTTCATATGCCCCTAGACAC
ATCTCTTTCATCTCAGCCATCAATTGATCCCATGTCTTTTGATCATGTGGT
15 CGACCACGATATACAAGTGATTAGCAAACATACCTAACATATTTTCAGTA
TCGCGATGAGTACGCGCACTGATTACACTACCAATAGCGATATCGTCCTGA
CGTGATATTTTGCAATAATACCATGATTGCCTAGCAAAGAACATAAAG
TCGTGCACTTGATGTTGTTCTACATAAGATTTCGAATTGCTGTTTGATTTGA
CGATTGTAATGAAACGTCAACATATTACCGTTGGTGTGTTTAATACTTGGT
20 CTAGGATAATCCGTAGGCATATTTAATATTGGAACCTGATTTTCAAATGTC
TGTAACCAAAAGTGACGTTGTTTAGATAAGTCTCTGTGCACCATCCACTCA
CTATAATCTTTTACTGAAGCTTAAGTTACAGGTAATGATTTATCTTGGTAT
AAAGCGTTCAAATCAGATAGTAAAATCGTGTTACTCATACCATCATTAATA
CTATGATGAGTATCCATAAATAAATAATCTTGTGTGGTCCATGTATATAT
25 TTAACCTCGCATCTGACTTGGTGTGTTCTAAATCAAACGGTTCATAAATGAT
TGAATAATATCTTGCTCGTTCGTTAGAGATGTCGTTACCTCTTCAAATCA
GGCGAAACATGTGTCGCAATACGTTGTTTAACTTCATTGTTCATCAATTACA
TATTGTGTTTCGTAATATTTTCATGACGTTCAATCAACTTAGATAATGCACGT
TGCAATTGCATAACATTAAGTTTCAGAAGATAATCTCCATAAGAATGGAATG
30 TTATACACTGTGCTTTTAGGATTGACCTTCCATAAAAGATACATACTTTT
TGAGACGCACCTTAAATCAATATTGATACGATTCAATTTGCTTTGGGAATCACT
TCATAGACATCATTTTGCAGTTCTTCAATTTGTTGTCCAAGTTGCTCTACA
GTAGGCGATTTCATTAAATCACCTACTTTAAGACGTTTTTTTAACTTTCT
TCAATACGGTTTACAACATATGTTGCTCTAAGAGAGTGTCCACCTAGTTCA
35 AAGAAATATCTTTAACACCTACCTGATCAACATGTAAAATCTCTTCGAAA
ATACGGCAAACGTGTCGTTCAATATCGTTACGTGGTCTACATAAATTTCTA
TTATTCTTTAGATTAAATTCAGGTAATGCACGCACATCTAATTTCCCATTC
ATCGTGATAGGTATACGATCCACCTTCATAAAATGCACAGGTATCATATAT
TCAGGTAATGTTTCACTTAAATATCTTTTAAATGACCTGTTGATTTAAT
40 TGCGATGCTTCATAATATGCCACTATTTGTTTATCTTGGTCTTGCTCTCGA
ACGATGACTACAGCTTTTATTAATATCACGTATAGCTTCTAATGCTTTTCA
ATTTCTGATAATCTATTTCTAAAACCGCGTATTTTAACTTGCTTATCGATA
CGACTAATAATAATCAATATAGCCATCTTCTTGAAGACGAACATAATCACCG
CTTCGATAAAGCATTTTCATTGTTAAAAGGTGACTGAATAAAACGTTACGCA
45 GTAAGTTTAGGTTGATTTAAATAACCTTTTGCTAAACCTGCACCACCAATG
CACAATTCACCTGGAACACCTACGCCACAAATACGATTACCTTGCATGACA
TAACTGTGCTTCCACTAATAGGTAAACCAATAGGTATACGTGAAGGCATC
TCTTGTGGAATCGCAAAAGTTGTAGTAAATGTTGTATTCTCTGTTCGGTCCA
TAACCATTGATTATTTGAGGATGACACTCACGCGAATTTAATAAGTGAACC
50 CATTTAGCATTTAACACTTCCCCACCAATAAGTAAATAAGTTAAAGATTCT
AGTGCTTCGATTCGCTAGCAATTTGATTAAATAAAGATGACGTTAAC
CACATCGTGTTGACTTTTATTTTTCAGTAATAGCTTGATCTAACAATTGAGGA
TTTAAACAACGTATCTTTAGATGTAATGACTAACCGTCCACCATTCAATAAT
GGACCATATATTTCAAAGGTTGCTGCATCAAAAGCTACTGTTCTTGATAAT
55 AAGACGGTTGTATTTTCGTTCAATTCGACATAATTTGGATTGTGTACTAAG
CGATCAATTCCCTCTATGTGGCACCAGTGTCCCTTTAGGTTTACCAGTCGTT
CCTGATGTATAGATGACATAAGCGATATCTTCTGAACAATTAATGCCCTCTC
GGGTTATCAATATCATGTTCTCTTGAATCAACTATCAATTCATATCCATT
TGAGGTAAACCTGATTGAAATGATGTACGATATGTTACAACCGCTTAGGT

TTTGCGTCCTCAATAATATAATTCATTCTTTCTTCAGGATAATCCGGATCA
 ATTTGGTATGTAGCCTGCACCAGCTTTCAAGATCCCTAACATTCCAATAATC
 ATTTCTAAGCTGCGTTCTGCTATTAATGCCACCATATCATTAGGTTCAATC
 TGATGATTTAAACGCAATCTATATGCTAAAAGATTTCGCACATTGATTTAAT
 5 TGATGATATGTCATCGATCGTTGTTTCATATTGCAACGCAATACTATCGGGA
 TGTTGGTGCACCTTGACGTTTCAACACCGGTTTTATAAGTGACC
 TCACTATAATTAATGTCATTGTTTGAAGATTGATGTCATTATATTTAGCT
 AAATCTGAGCGATCACTGAGTTTATATCTTTAATTAATAATGATGGATTT
 TGTGTAATTTGCATATAAAATCCCATTAATTAATTTGACTAACGTATTTACA
 10 GTTAATTCATATAAACGTTGTCATTATAAACTATATCAAAACCACTGTC
 ACATGAGGATAAAAATCTAAATCCGCGAATGATGATGCTTCATTGTATAGA
 CGATGTATATGATGACATATATGATTAAGCTCAAATGTTCTTGTACTTTT
 TCAATATGAATCATCGTTTCTACATCTATTTGTACTGCTTTAGGTTGAACA
 ACAAAGACGCGACACACATTTGAAGCTCTTCAAGTAGCGCGGAACACTCA
 15 TCCACCATATCTTTTACGGCGTCACTTTGACTAATACTTAAATTAAGTGGG
 GCAATGTCGTGTTTAAACCATCATATTCTCAGTATTTTATTATCTATC
 GAAAAATGATGCTTAAACGTCACATCAGAAGACTGACTTATAAAATGATGC
 GCTAAGTAAATACTTACAATCCACGCATCAATGTCTAAAGATGGTAGTTCT
 GCATGCAAAGAACACATTTTCTTTTTCAGATGATTCCTTTACAGGTATG
 20 TAAGCGTAATTATCACCGTCTGCATTCAATAC

Sequence 3431

step.1003a07.cons.ok

AACATTAAGATACTCCCTTGCTCATTTACACCCAACTAGTGATAAATTGG
 25 TATTATGTCGCGTGATAACATTTTCTAATAGCATCGTGTTGATTCTTT
 AACTTTACCAACGTAACCATAGATATTTTATTTAAAAATTCGGCGCTTAC
 AAGTTCTTTTATCTTTTCTTTAAATGAAAAATTTCTAAATTAGATTCTTT
 AAGTCTTTCTAAAAGTTGATGGCCAACTATGTTAATTAAAGCATCATATAT
 30 TAGGTTAAATCATTTTGTAGCAGTTACTCTCATGCCATCAATTTAATAGT
 TGAATGATTGTTTTAGTTAATAAATTAAGTATGACTTGGCCGCCACCGTG
 AACAATAATTATTTTTTGTTTTCAAGGTGCCAAGCATTAATTTGTGTAT
 AAATGCGTCGTTTAAATTTTCTATAGCTATACCAACGAGCTTAATTACGAT
 GATGTTTTTCATATGTTCCCTAAGTGGTGATAGTGCGTTGATTTTAAAC
 35 GTAGTCATACGATAAATCACATCCCCATGCTTGACCTTCACAATTCCCTTG
 ATGAAGGTCAAGCTGAATTGATATTTCTTCAGCACTCATTATTTCTTGAAT
 TTCTTCTTTATCGTACTTTACTGGTGAGGATCTTATTAATACAGGTATCCT
 ACCTATAAAAAATGTCTACCTGATTAATATCAAAATATGTTTTAGCATAACC
 TGCAGCAGCAATAATTTTACCCCAATTAGGATCTTCGCCAAAAATGCGGT
 40 TTTTACTAAACTTGAACCCACCACACTTTTAGCAATCATTCCTGACGCACT
 AGATTCTTTTGCACCTTTAACCCTGACTTCTATTAATTTAGAAGCACCTTC
 GCCATCCCTTGCAATACTTTGTGCTAAATCGGTCATAATATATAGAAGCAT
 CTGCTTAAATTTATAGTAGTCTTCGCTGCTTTTTTAATTTTCGTTATTATT
 TGTACATCCATTTGACATCACAAGCACCATATCATTTGTTGAAGTGTACAC
 45 ATCTACAGTGATTTGATTGAATGTAACCTCAACCACATCTTTTAAAGCCTG
 TTGTAATCTTTGTGATGAGATGTTAGCGTCACAGGTTATAAATGCTAGCAT
 TGTAGCCAAATTAGGATGTATCATTCCTGACCTTTTGTACACCTGCCAT
 CGTTACTGTATCGCTACCAAAATCTTCGTTTACAACGCATGTTTTGTATG
 AGTATCCGTTGTTAATATCGCTTTTGCAAAGTCATCAGCATTACCGTTTTT
 50 AACTAGTTTGGAAAAGCCATTCTTTAGAATAGACATTGGCATCACCTTTCC
 AATAACACAGTAGATGCGACACCAACATATCTGGTTGAATTTGTAATTT
 ATTTGCGGCCAGTTGTTGCATTTTAAAGCATCTTTTCTCCTTGTTTACC
 AGTACAAGAATTAGCAATACCTGAATTAACAACAATAGCTTGCATTTTACC
 ACTTTTTTCGATGCTGTTTTTGTTAATTTTAAATGGTGCAGCAATGACCTT
 55 ATTAGTTGTAAATACACAGCTACATTTGCAGGTAATCTGAAACAATCCA
 ACCAAAGTCTAATTTTTTCTTTTAAAGCCAGCGTGCAGACCATCAGCTGA
 AAATCCAAGAGGACTTGCAATATTTCCCTTAATTATATTCAATTTAGTC
 CCCCATATAGATATATAAAGCGCCAAAAATTAACCATCAGTTTCGTGCG
 AAATTAACATCAGATTTCATGTTTTGAATTGCTTGTCCAGCGGCACCTTTA

ATTAAATTATCAATAACTGATGAAATTGTTAAAACACCCGTGGTTTCGTTA
TAAACAAAACCAATATCAGTGTAAATTTGTTCCATTACTTCATTTAACTGT
GGTAATCCATCTTTAATTCTTATGAAAGGCTTATTTTATATACATCCTTG
TAGGTAGATCTATTTGATTATTTTAACTCCGTTTCAAGCCGTGTGTAT
5 ATAGTTGCAACAATTAATAGAACTAGTGATTGATGTCAGATTGTACCT
TAGCAAGTGCTTTTTCAATCTTATCACTACCTTGTAGTGAACGTGTAATTG
CTTGATATATCTGTTTTTCATCTGTGACATGTTGATCTTTTGAATCCCTAA
TTTGATCAGTATCTGATGCTATACGTTCTAATGTTGAGCGACGACCATCAT
TACTTTTATCGTCTTTACTATCACTATCGGATTGATTGTTGTCGTACTAG
10 ATTTCTCTTTGTTATTTTCGTTTATTACGATGGTGATGCTTGCTGATCG
CTTTGTTCAAAAAGCTTATTTATAGAGTCAGACTTAAAAATCTTGTGATT
TTATCTAATTTATCAATAACCGACTCTTCAGAATCATCAGTAGTTGCATTG
CGTGTACTAGCATTTGTATCGTGACTTTTATTATTTGAATCATCTTTTCA
TCGGTTGTTGAGTCACTAGGTTTAGAAGCACTATTTTCACCTGTTGACGG
15 TTATCAACATTTTAGCAATTGAATCTAAATCCGACAGTATTTTATCTAAG
TCTTTAGTCATATCATCATTTATCTTTTTCAGATGATCGTCATCTTCAGTT
TTTGATGATGTGGCACTAGAGTCTGTCGATTTTAAATGATTATTTTATCT
TTTGATCATCATTTATAATACGGTCAATACTTGATTGATTATCGACTTTATCT
AAACGTTCTTCAAACTTTTGATATCTGCGTCAGAAAGTTCTTTCAAACTCT
20 TTAATATCATCTTTAGCTTGTGATTGTTGTCAGCGATGGAGTTACTATTA
GAAGTATTTCAACATTTGATGATGTACTTTCTGAATCTGCAAAAGCATCT
TTTGTGAAGCCATAACCAACGAGCAACCACTAAGTAACAACGCAGAACAT
GACAGCTTGAGGTAGACACCAATTTTATCTGTTCTCTTCATTTACACATAC
TCCTTTAAATAAAGTTACCTGTATTATTCTTTAAATAGATTATATGATTTA
25 ATCAAAAATATACATATCACATTATAATACATAATTTTCATAGATTGTAGA
TAATGAACCTAAATAACAAATTTAAAAAACAAATTTCTTGTAATTATTAG
TAATTTGGTATTACACTTCTTT

30 Sequence 3432

step.1003a08.cons.ok

TGGATGAACCTGTTGTCTTTAATAATATGACTTGCTTCAATAAGATCAGA
TAATCTTGCATTCGTGCATGAACCTAAAAAACATAACCTAATTTTATATC
TTCCGGCTTTTGACCTGGGTGAAGTCCCATATAGTCATAAGCACGTTGGTC
35 ATTTGCATTTTAAATTTCTGGGAATGGATTACTAAAACTAACTCCCATTTT
TGGGTTAGTTCCTCAAGTTACTTGAGGTTCTAAATTTGTTACATCAAGTTC
AATAACTTTATCAAAATAGGCATCATCATCAGAAATAAAGTTCTTTCCACCA
CGCCATAGAACATATCAAAATCAGTAGCATAAGGACGACCTTTTACGTAATT
AAAGGTTGTTTCATCAGGTTGCATTAAACCATACTTTGCTCCTGCTTCAAT
40 AGCCATATTACAAATCGTCATACGTGCTTCCATGGATAAATTTTATAGT
TTCTCCTGTAAATTCAGAGCATAACCAAGTACCGAAATCTACACCATATTG
ATTGATTAAATATAAAATAATGTCTTTAGCATAACTCCAGCTGGTAAAGA
ACCATTTATATTAATTTTCAAATTTTAGGTTTGTGTTGCCATAATGTTTG
TGTGGCAATAACATGTTCAACCTCGCTTGTTCCTATCCCAAATGCGATAGC
45 ACCAAATGCGCCATGTGTAGCAGTATGTGAATCTCCACATACAATAGTTTT
TCCTGGTTGAGTTAATCCAGTTTCTGGTCCAACCATATGCACAATACCTTG
TTCGTCAGAACCCATATCAAAAATATGTACACCAAAGTCTTAGCATTTTG
TTGTAAAGTTGTAATTTGTTTATTAGCAATTTTCATCTTTTATATTAATAAT
ATCAATTTGTGGGAACGTTATGATCTAAAGTTGCAAAGGTTAGATCAGGTCT
50 TCTGAGTTTACGATTTTGTATTCTAAGTCTTCAAACGCTTGAGGAGAAGT
GACTTTCATGAATGAGATGTAATCAATGTATAATAATTTGTTGGTTACCTTC
TTTTCCATGAAGCACATGTTTTTCCATACTTTATCAAAACAGTGTGTTGACC
CATAATTTTCTCCTCCTTATAGATATTTTCTTTAAGCAATTTAAAAAT
TTCTGAAGTTGATATTGTTCCACCTAAATCTGCAGTTGTCTTATTAGATTG
55 AATAAACGAGTAAACAATTGACTCAAGTTTCGTTAGCAGCATCATTTGATT
TAAACTTTCTCTTAAGCAAAGTGCTAAAGATAGAACCATAACCAATGGATT
CGCTTTATCTTCATTAGCTATATATAAAGAACAATGTGAACCTGACCTGTTA
AAAGTCGTGCAACAAGTCTCCACCTAAGTTGTGGATAGCTTTTACATTTT
CATCATTTGGTAGCATTTGATAGTTGTACCATATGGAGCGCCATAGTCAATAC

CAAAGTGAGCACCACCATTAAACTATAACCAGGAGCACCACCATTAGGAT
 AATACCCCGTAGTAATTGGGAATTTAGTGAACGAAGAACCATCTCCACCGC
 CTGCAGAGTCTAAACCTTCATTAAAGAAGTCTTTAATACCATTTTAAAGTT
 TCTTAATCATAGCTGTCAATTAATCAAATGGTAATGGTGCATTTTTCATAA
 5 AGCTTAGGTTAAAACCAACTTTCCTCAAAGACTTTATTTACCAATTTTCATAG
 GATGACCTACATAATCAAAAACATCTCCAATACCCTTGATAACAGTGCTAG
 ACACATCTTTAGCTTTATCAAGTGCTTTTGTACCTATGTCTTTTGTCTTTGT
 CAAACGTAGTAGCACCTATGTCTCTTGCTTTACCACCAATGTCTTTTACTT
 TATTAACATCATTTTACACCAGAGCTAATAACATCTAATATACCGTCTT
 10 TCTTTTGTAGTACCTTTACTAAATTTAGGTAGTTTTTTCTTCTTAGTGTCTAT
 ATCCTGAATTAGATAAAAATAGCATGTGTTTGCAGCCACTATATACACTTG
 AGCCTTTAGGTAAAGAATGTTGTGTATCTTTGTAGGTGTAGTGCCATTCT
 TTCCGTTAGGATAACGGATCATTTCATTTCTAAAGCCACCAGGTCCATTAC
 CTCGACCTCTATCTCTACAGTAGCCATTGTACCTTTATTGATTTTACCGT
 15 TGGTAATATAACTTTGTGTATGCGTAGATTGAGTACCTGTAGATAATTTAA
 TCTTAGGTATTTATCCATACCTAATTTATCTGCGACCCAGTTAACACCTT
 TGATTAATCCATTAAGGCCTTTTAAACGCCTTTAACCATTCCACCAAAAA
 GATTCCTCAATTTTACCGGTAACCTGTTTGTATACCGTCTCTCATTTTATTCA
 TGGTGCCCATAACTTTTCGATTTCATGCCATTTACTATAGATACAGTAGTAC
 20 TTTTAATTCATTCATTTCTTGCTCATGAAGCCACCTACAGCATTCATGG
 TGTATAGTAGTACCTTTTTTAAGTGATCCCCAAGCACCTTTGACGCCTGACC
 ATAGAGCTTTTCGCTTTATTAACAGTACCTTTTTTAACTACTGTTCCATTTAG
 AACTCATGAACTGCCAAGCTTTAAATATGCCAATTGTACCTTTTTTGA
 GTGCATTCCATGTATTTTTAACTCCAGACCATAATGCTTTTGCTTTATTCA
 25 CTACTGATTTTTTATAGCCGTCCAGATTTTAACTGCGGCATTCTTTACAG
 CATTAAATATTACAACAATACCTTTTTTAAATGCGTTAAATACAGATAGAA
 CACCTTTGCGCAAAGCTCGAACGATTCCCTAACACACCATTTTTTAAATGCAG
 TCCAAACTTTAATAGAGAACTCTTAATAGCATTAATATCGTAACCTACTA
 TGCGCTTAATAAGGTTGATATTAATTTTACTTGCGCAACATATGCTTTAA
 30 TAAACTAAGTAAATTAGAAGTACAAGATCGAACACAAGCAGTAATATATGC
 GTTCCAGCATAATTTAATTCAATAAATTATTAAAGGCGAAAGTAAAGATAC
 ATCTATCGATACTTTTCGCCTTTCAATATTATCAGTCGCTTGTGTGATGTTT
 TTC

35

Sequence 3433

step.1003all.cons.ok

TTTAATTGTTCTACTTCTGACAGCTCTTCTAATCCTTTACCGTAGGTATAT
 TGTTTACCAACTTGTTGTGAAAATCTATAACTTTCCCCATTTCGATACCAT
 40 CGCCACCAAGTTTTCACCTTGCGTCCATTTCTAATATTTAATTCTTTTATA
 ATTTCTTTTGTGAAAATCCTGCTGCTTTTCATTTCAACTGCTTTTACTTTT
 GTTTCTACTGAATAAGAAGCTCTTTTCATAGAAAAAAACCTCCGTATGATT
 CATTTTAAATATGAATTCACGAAAGTGTTTTTATATAATTTCCACAAATTG
 GGGTCAGTCTAATAATGAATTTTCGTGATTGTAGGCTTTTTATTTTATATAT
 45 TCATTAATTTCTCTTCTTGACGGATACGTACAATTAAAAAATAAGCATATG
 GAACATAAAGAGTGTTGTGTATGTAGCATTTGTTAGTAGTAATACACCAA
 TTAATTCGGGAATGATATTTAAAAAATAGTTTGGATGTTTCGTTACTTTAT
 ATAACCTGACTTAATAATAGGGTGATTGGGCAGTATAAATAATTCAATG
 TCCAAATACGACCTAATGTCTTAATAACTATAAATAGCATGATATAAGCGA
 50 TGATCAATATGATTAAAGCCAATGCCGTTAAGTAGACTAAATGTATCCTTAC
 GAATGAATGCTTCTATAGCTGCACTCATGTAAATTAATGATGCGTAATGG
 CTAGATATTTTGAATTTTTCACACCATATTCCACCGCGCCTCTACCTTTA
 GCTGTTTTGCGTGTTGCATAGATATCTTTAAGCTGATGAGTCGAATACAGA
 AAAAGATAAATAAATAGTTAAATCATGATGTCTCCATTGTTTGTATATA
 55 TATATATTTCTTTTGCTCGTACAATAACAAAGTATTTTATATAGAAAAATGTTT
 CAAATGCCATTAACATTTTATTAACCTTTAGCACCAACTCATAACTCTTCTA
 ACACCTCACATATTAAGTTCAGTTTTTCGGATAAATTAATACTTTTAAGGA
 TATTAAGCGCTTACATTGATGTGATATATTTTTTTTAAACGAAGATGATAGG
 GGTGTTAGGATGAACTTTTTTGATATTCATAAAATGCCAAACAAAGGGAT

5
10
15
20
25
30
35
40
45
50
55

ACCATTAGCTGTACAACGCAAAATTATGGCTCAGAACTTTATGCAAGCGTT
TTTTGTCTGTTCTTTGTTTACATGGCGATGTATTTAATTCGAAACAATTT
TAAAGCGGCACAACCGTTATTTAAAGAAGAAATCGGATTAACAACATTAGA
ACTAGGTTATATAGGATTAGCGTTTAGTATTACTTACGGTTTAGGAAAAAC
AATACTCGGTTATTTTCGTTGATGGGCGTAATACGAAACGTATTATTTCCCTT
CTTATTAATATTATCTGCGATTACAGTACTTATTATGGGATTTGTATTAAAG
TTATTTTCGGTTCTGTGATGGGGCTATTAATTGTATTGTGGGGCTTAACGG
TATATTTCAATCTGTGGGTGGGCCCTGCAAGTTACTCAACGATTTCAGGTG
GGCGCCTCGAACAAGCGCGTTCGTTATTTAGGCTTTTGGAATACATCACA
TAACATTGGTGGTGTCTATTGCTGGTGGTGTGCGACTTTGGGGCGCGAATAC
ATTTTTCACCGGTAATGTGGTTGGAATGTTTATTTTTCCTTCCGTCATCGC
TTTAATCATTTGGGATTGTGACATTATTTATTGGTAAAGATGATCCAGAGGA
ATTAGGTTGGAATCGTGCCGAAGAAATTTGGGAAGAGCCTATCGACCAAGA
AAACATTGATTCTCAAGGTATGACTAAATGGGATATCTTTAAAAAATATAT
CCTTGGAAATCCTGTGATTGTTGGATTGTTGTGTATCTCTAATGTTTTGTATA
TATCGTCCGATTGGTATTGATAACTGGGCACCGCTATACGTATCAGAGCA
TTTACATTTTAATAAAGGTGATGCGGTGAATACTATTTTACTTTGAAAT
AGGTGCATTAGTAGCTAGTTTATTTGTTGGGGCTATATCTCAGATTTATTTAAA
AGGTCGTCGTGCGATTGTAGCGATTGGATGTATGTTTATGATCACCTTTGT
TGTACTCTTTTATACCAATGCAACAAGCGTGACAATGGTCAATATTTCTCT
ATTTGCATTAGGCGCTTTAATCTTCGGTCCACAGTTACTCATTGGTGTATC
TCTGACTGCTTTGTTCCATAAAATGCAATTAGTGTGCTAACGGTATGAC
AGGTTCAATTTGCATATCTATTCGGGGATTCAATGGCTAAAGTGGGTCTGGC
TGCAATCGCTGATCCAACACGTAATGGTTTAAATATTTTGGGTATACGTT
GAGTGGTTGGACAGATGCTTTTATGTATTCTATGTAGCTTTATCTTAGG
AATGATATTATTAGCCATTGTTGCTTATTACGAAGAAAAGAAAATTAGAAA
ATTTAAAAATTTAAGATTAAGTGAATTTAAATATATACTCCCTACTATAAAT
TTTATCATATTAGGTAGGGAGCTTTTTCGTGATTGATTGCACTATTTTGAT
TGCTTATTATAACTACTTGGTGACATATGTAAATATTTTAAAAATGATAG
30 CAAAACATTTTATACTCAGAAAAACCTACTTTTTCAGCAATTTTATAATGC
TTGTAGTGCTGGTCTAAAAGATGTAATGATTTTAAATACGATAACGATT
AGATAATCAACTATCGTAATGCCTACATGTTCTTTAAACGTCCTCATGGCG
TATGACTCACTTACATCAATAGGATTAATTAAGTCAAGAACAGTCACTTTC
TTGTGATAATGTTGCTTGATTGAGACAAAATTTGATTGACATAATAGTCA
35 TCGTAATCAATTTTAAATAATGGTTGAAAGGCAGTATGATATGCCGCGTCA
TCATTGGTAGAATGTGGGCGTTCTAATAACCTTTGAACTAATATGTC TAGA
ATATGCTCTAAGTGTGGTCTACTGGCTTTAGTAAATAATCAAGAAC
TGATGTTGTATACCGGCTTTTCATATATTTCGAAGTCGTCATAACTCGATAAA
ATGATAATCTGGCAATCAAGGTCCTTTGATGTCATCTAGTAGGTCAACACCA
40 TTTTTCGAGGCATTTCGTATATCAGTAATGACGAGTTCAGGTTGATGTTGG
CGAATTAAAGACAATGCTTCCACGCGCTTTTGGCAGTATAAACAGTGGTG
AAATGATAGTCTCCCATGGAACCATTTGCTTTAAGCCTTCTCTTATAATC
CTTTCATCATCACAATAACTACTTTAAACATCTTTATTCCTCCTAGACAA
TGGGTATTTGGTAACACATTAATGTCCTTGTGGCTTCTTGAGAAAATGT
45 GCAGACGTGCATATGTTCCATATTGAATCATGGCTCTATTATGTAAATGAT
TTAAACCTAGGTGTGTCGTATCAAAAAACATCGTGATGAAGGGATTGGCGCA
CGTGTCTAAATGTGATGGAGACATACCGATGCCATTATCATGAACCAGAA
TATGTAATTTGCGCTTCGTAAAGTCTGATACGAATTGTTATCTTTAAAGGTT
CACTATCACGACCATGCTTGATGGCATTCTTACGAGTGGTTGAAGCATCA
50 TCTTACCAATTGTTTGATGTTGTACACCTCAGTAGCATCGATGTAAAGCT
GTATCATATCATCGAAGCGGATATTTGTATGGCAACATACTGCTCAATGT
AGCTTAATTCCTGCTAATTTGACTGTGTGCGATGCTGTGCGTAGAGAA
AACGTAGCATTTGTGATAGTTGTTGTATGACTGTTGTGCTACTTTGGGTG
AAAGAGGAAT

Sequence 3434

step.1003b04.cons.ok

GCGCGCAATTGAACAAAAGAAAAAGGTATTGAAACATTAGTAATTGAAA

AAGGTAATGTTGTTGAATCAATCTATAATTATCCAACACATCAGACTTTTT
TCTCATCAAGTGATAAATTAAGTATCGGCGATATTCCTTTTATTGTTGAAG
ATAGTAAGCCAAGACGTAATCAAGCGCTTGTATATTATAGGGAAGTCGTTA
AACATCATCAACTTAACATACATCCATTTCGAAGAAGTTTTAACAGTTAAAA
5 AAATAAAACAATAAATTTGCAATTACAACACAAAAGGTGTATATGAATGTA
AATATTTAACTGTTGCTACGGGTATTATGGTCAACATAACACTTTAGAAG
CGGAAGGGGCAGAAATTACCAAAAGTATTCCATTACTTTAAAGAAGCACATC
CGTATTTTAATCAAAATGTTGTTATTATTGGAGGCCAAAACCTCGCTGTTG
ATGCTGCCTTAGAATTAGAAAAAGCTGGTGCATTAAGTAACTGTTTATATC
10 GTGGCGAACAGTACCCTAAAGCAATTAAACCATGGATATTACCCAATTTTCG
AATCATTTCAACTACGAAAAAATTACGATGGAATTTAATGCGACAGTAA
CCAAAATTACCGATCATTCAGTGAATTTATGAAAAAGATGGTCAACTTATAG
AAATTGATAATGACTACGTTTTTGTCTATGATTGGTTATCATCCAGATTACG
ATTTCTTAAAAACAATAGGTATTGATATCCATACCAATGAATATGGAAGT
15 CTCTCTTTTATAATCGAGAAAACATTCGAAACAAACGTCGAAAATTGTTATA
TAGCTGGTGTATTGCTGCGGGTAATGATGCAAACTACTATTTTATCGAAA
ATGGTAAATATCATGGTGGTGTCTATTACACAAAGCATTTTGACAAAAAAC
AAACACCTCTTGAACATAGAGCTTACAAAAAGCTTGGGATACTAAACAAT
GTCCCAAGCTTTCTTAATGTATAGTCAATCAATGACTCAATTAATATGCAC
20 ATATTAATAATAGGTAATCCTTAATTATAGTATAATTGCACATAAAGAAGTT
TCATAGACAATGCATACAACTAATAGAAATCGGAGTGTGACAAATGAAGT
ATACGAATTTTGTCCCACTCTTTTAAATTACTCTTCGAAATATCGCTCTAA
TTGATTTTGAGTCATGTCTTGACTCAAACCACTAGTAATTTTAGCCTTGC
CTTTGGTCCATTTAAACCGTTTCGAAAAAATAACACCATTAATTTTCAAATC
25 TCGGCCACCACCTTCATAAGCATATACAGGACTAACAATACCATTGAATGA
TCTAGATACGAGAACTAGAGGAATGTTTTCTTTAGACATTGCTGTAGTCC
ATTAAGACAACCTTTTTGGAAGGTTACCTTGTCTTAGTGCTTCGATGACTAT
ACCATCAACGTGTTGTTGTGAATAAAATGATAGTACATCATCTTCCATACC
CATGTATGCTTTTACAAGTGGAACACGTAAATTTACATCGATATATTGGTA
30 GGTAGTTTGTCTGTAAGGATGATGATAAAATTGTACTCGATTCTTGGTAAG
TACACCTAGAGGCCCCGATTAGGACTTTGAAATGTATTAATATTGGAAGT
ATGTGTCTTTGTACATTACGAGCAGTGTGAATCTCATCTAAATACGAC
CATAACACCTTTATGATTAGCCTCAGATGAAGAAGCAACCTTTATAGCAGA
AATAAAATTATAGAGACCATCGGAACCAATTTCAATTGGATGATCTCATTGC
35 TCCAGTAATTAATATAGGCTCTTGAATATCAATTAATAAATCTATTAAAAA
AGCTGTTTCCCTCAAGTGATCTGTTCCATGAGTAATGACAAATCCATCATA
TATATTTCTTTAGAAATATGTAATGATTTCGTCTCTTAATCGCACAAACATT
CGAAATTTGTCTATATGCGGCGAGGGTATATTTAAAGATTGATTTCGTCAAC
CTCTGCATATTGACTAATGATATTTTGATGTTGTGATATTGGATTTCCTTC
40 ATTCGTTATCACTTTTATTAGTTTGATCTTGTGACATACTTATTGTGCCACC
AGTATGTATGATAAGTAGACGTTTCATTACGAATTCCTCCTATTATTCTT
TACTTTTTTATGATAAAATAATATCTATAAAACAACAAGGAAGGTTATCAA
ATAATGAGTTCAATTAATATTGCACTAGATGGCCAGCTGCTGCAGGTAAG
AGTACAAATTGCTAAACGTGTAGCCAGTCGTCTATCAATGATATATGTTGAT
45 ACAGGAGCAATGTATCGTGCCATTACATATAAAATATTTACAAAATGGCAAA
CCCCGAAAATTTTGATTATCTGATTAAATAACACTAAACTTGAGCTTACTTAT
GATGAAGTAAAAGGGCAAAGAATCTTACTAGATAATCAAGACGTCACTGAT
TATTTAAGAGAAAATGATGTAACACATCACGTATCTTATGTTGCATCTAAA
GAACCAAGTGCGTTCAATTTGCAGTGAAAAATACAAAAGAATTAGCTGCTAAA
50 AAAGGTATCGTTATGGATGGCCGAGATATTGGTACAGTTGTATTACCAGAT
GCCGAATTAAGTTTATATGATTGCATCTGTTGCTGAACGTGCTGAACGT
CGACAAAAAGAGAATTAGCAACGTGGCATTGAATCAAAATTTAGAACAATTA
AAGGAGGAAAATTGAAGCACGAGATCATTATGATATGAATCGTGAATTTTCG
CCATTACAAAAAGCCGAAGATGCTATTACACTTGATACAACCTGGCAAATCT
55 ATAGAAGAGGTAACAAATGAAATATTATCTCTACTTTAAATGTTAAATTA
AACTTTTTTATATTGTTTATTGTAAACATTTAAGCGGAGTAGAATTAATAAT
CAATGACTCGAAAATATTGAATATATTAGTATTGATTTATAGAAAAGCCAC
GGACGCTCTGTCTCTGGCTTTTATTTTGGACCATGAGAACTATGTAAATA
AACTTATTTTAAACCAAGTCCCCCTCTATAGAGAAGTAATACTATGTTTATT

5 ACCAATTTTAATATATCTTTTTCATTTTATTCCTCATATTATTATGAGCT
 TAGGGTTAATTGCCATAAAATTTATGTTACAATAAAAAATAACAATGTTATT
 AAATATACTACAACTAAATTGACCATCACTTAAGTGGTGCTCTGTTAAA
 CTTTTGTTATAAATTAAGATAGAAGCAACCTATTCATATAAAGTGGTATC
 10 TCTTGACAATTTTACTCGTTTATAAGATGTTATAATTATGTAGTGTTAAG
 GAGGCATACAAGATGACTGAAGAATTCAATGAATCAATGATTAATGATATT
 AAAGAAGGTGACAAAGTCACTGTGAAGTTCAACAAGTAGAGGATAAACAA
 GTTGTTGTGCATATTAATGGTGGCAAATTTAATGGAATTATTCCTATTAGC
 CAGCTTTCAACACATCATATCGAAAAACCCTAGTGAAGTTGTAAAAGTCGGT
 15 GATGAAGTCGAAGCATATGTCCTAAATCGAGTTTCGACGAAGAAAAATGAT
 ACTGGGGCATACATTTTATCAAAAAGACAAC'TTGAACTGAAAAATCTTAT
 GAATATTTACAAGAAAACTAGATAACGATGAAGTGATTGAAGCTGAAGTT
 ACTGAAGTAGTTAAAGGTGGTTTAGTCGTTGACGTTGGTCAAAGAGGGTTT
 GTACCTGCTTCTCTAATTTCAACTGATTTTCATTGAAGATTTTCTGTATTC
 20 GATGGTCAACAAATCCGTATTTAAAGTGAAGAAGCTTGATCCTGAAAAAAT
 AGAGTCATTTTAAAGCCGTAAAGCTGTGGAACAGTTAGAAAACGACGCTAAA
 AAAGCTTCAATATTAGATTCTTTAAATGAAGGCGATGTTATTGATGGTAAA
 GTTGCTCGATTAACTTAAC'TTGGTGCTTTCATTGATATTGGTGGCGTAGAT
 GGTTTAGTTCACGTTTCTGAATTATCTCATGAACATGTTCAA

Sequence 3435

step.1003b05.cons.ok

25 ATCTACTATTTTCTCAACATCTTCTTTTGTATGAATGCCACTCTCTGAAAT
 GTAGCAGCAATTAGACTTTTTAACTTAAGTAATTTATTTGTATGTAGAAC
 ATCGGTTTCAAATCGTTTTAAATCACGATTATTAACACCAATAATTTTAGG
 GTTAATTTGGTGTGCACGTTCAAGTTCTCTAATTGTATGAAC'TTCTACTAG
 AGCTTCTAAAT'TATGGTTTGTTCGATATGAATACAATTC'TTTTAAATGGTC
 ATCACTTAAATATTTACTATTTAATAAAATAATAGATGCACCAGCTCGTTT
 30 TGCAACATCTATTTGAATTTTATCAATAATAAAATCTTTACATAAAACTGG
 TAACGATGTTATCTTTGATAACTGATTTAATCGTTCAAAACTACCGCCAAA
 GTATTTTTTCATCAGTTAATATTGAAATAGCATTAGCACCATATTTTTGATA
 ATCTTTAACTTGTGAACAAGATCACGTTGCGGTAATTGAGGTACAGATGG
 GCTTTTTCGATTTTATTTTCAGCAATAACTGATAATGTTCTATCATAGTTAAG
 35 TGAATCAATCAGCTTCTCCTCTTAACATTTCCGTTATCTTGTAATTTTC
 AAGTTTTTTATCATAGTATTTACGCTCAAGCAAAGTTTTTTTATACTCAAT
 AATTTCAATTTAAATAGTCATTACTTAACCTCCCATTTTAAATATTGTTT
 CATTGCCATACCTGTATCAATGAGATATTTGCTCTCTCTACGCCATGTTT
 GATACTTTCCACTTGCTCAGCAACATATAAAGCAATTCAGCATTTAACAA
 40 AACTACATCTCGTTTACTTGAGTGATCCGTGCCACTTAGGATATTCAATGC
 AATTTGTTTATTTGTTTGAAGTGAACCACCTATCAACGTGTCAATTATTGC
 ATAAGCTAAACCGACTTCTTCTGCTTTTAAACTATATTTTTTTAATGCTCT
 TTCGCTGCTAATTCATAAATGATATTTTCACCAGAAAGCGTGGCCTCATC
 CATCCCATTTGCACCATGAATTTAAATTTGCTCGTTTCTACCTAAATCCTT
 45 TAATGTTTGTGCTATATTTTCAAGTTGTGAAGCTTCATATAACCCCATCAC
 TTGATAAGTTAATTTGAAAGGATTAATTAATGGTCCAATCAAGTTAAAAAT
 TGTAGGTGTTGCAATCGATTTTCTAATTGATTGAAGCTTTTTCATCATTGG
 ATAAGAATCAGTTGCACTTATGAATGCTAATCC'TTCAAATTTAATGTTG
 CTCTACTTCGTTCAATTTGTTTGTATTTTATATTCATTTTCAATGTAATACATC
 50 TGTAATCTCTGAATGTGAAGTAATACTTTTATTACCGTGTTTAATGACTGG
 CACTCCTGCAC'TTGCATACAACAAAAGCTACAGTTGTAGAAATATTAAAGCT
 ATTTGATTGATCTCCACCTGTGCCACAAACACACATAGCTTTATTATAAAA
 TGGTTGGTTTGGATAGTTTGTCTGGATAAAATATTCAACTAAATACGTTAG
 CTCATATTGACCATGTCTTTATTTGTATAAGCTTTCAATAATTCAACCTT
 55 TACATTGGTTTCTATATTTGAATCAAACAGTGTAAACAATAAATGATTGCAT
 ATCTTTTTTAGATAAAGATTTATTTGTTTAAATTTCTCAAGAAGGGTCAT
 TTTCAACTTCTCTCGTTGCTATATTTATAAAATTAATAATGATTGTTTC
 CCATTTTCTGTAGCGAATGACTCTGGATGAACTGTATGCCATAATGTTT
 CTGTTTTTGTGTTGAAAGATTGTATACTATCATATGTTTGGCCAGTGATT

AATAAACTTTCTGGAAATGAGCATGGATCACTAATTAATGAATGATAACGC
 ATAACCTTTGAAACTTGACTTAATTCCTTTGTATAGTGGTGACTGATTAACA
 ATAGTTAAAGTATCTACTTTTCCATGTTTGACACAACCTACCTGTATAACA
 TCGCCACCGTAATAACAAGTGAGCGCTTGTGCACCTAAGCAAACACCTAAA
 5 ATAGGTAAATTTTCGAAATTTTCAATTAATTTAATTAATAATTGATTGTCT
 AATGGATGTCCAGGTCCTGGTGAATAATAACTGCTGTAATACCTTTCAGA
 TTACATATATCTGTATCATCGGGATATTTAACAATCGTTTCACATTGTTTT
 TCTACAATATCTACAAGATTGTAGGTAAAAGAATCATAGTTATCTATTATT
 AATATCATGGAGTTACCTCCAATAAACTTTTAGCTTTAAGTTTTGTTCTT
 10 CAAGTTCTTTCTCTGGAATAGAATCATATACTACTCCACATCCTGCCTCGA
 CACTGACTTTTCTCATCGATAATCATGGTACGTATAGCCAATGCAAAAT
 CTAAATGATGATTACAGTTGATATACCCAACACCACCGCTATAGATACCTC
 TTTTATAAGGATAAGATTCTGTATATCTCTGTATAGCTCTAAGTTTAGGTG
 CACCTGAGACAGTACCCGTTGGTAGCAAACCTTGCGATGACGCTCATAGGAG
 15 ATAGATGGGGTTTAAATTCCTCCAATAAECTTCACTAACGATATGCATGACAT
 GTTCATAACGTTCTATTGTCTATTAGTTTGGTAATTTGTGAAGTGCCGTGTTT
 TACTTATTCGATGAATATCATTTCTTCTTAAATCTACGAGCATACGATGTT
 CACTCAATTCCTTTTCATCTTTTCATTAATGTCTTTTCATTATTTTCATCTT
 CTTTTTTATTTTGACCTCTTTTAAATTGTTCCAGCTATAGGATTCTGTATAAA
 20 CTTTTCCATCTTTTACCTTTACAAAACCTTTCAGGAGAACTTCCTATTACAA
 TCGGTACATCTTTATTAATATAATACATATATGGACTAGGATTTTGTGCGCT
 TTAAATTTTGATATAACTGAAAAGTTAATTGATGTAAATTTGTGTTGAAAAT
 GGTGTTTATAACTATAAAATTCCTGAAGGAACTACTTGAAACATATCTCCTT
 CAGTAATTTCTTTTTTAAATTTCTAATAGTTTGAACAAATTTGTGCTCTG
 25 ATATATTGGTGGTTATATGTCGAGGGAT

Sequence 3436

step.1003b06.cons.ok

30 ACAATTATATTTGATATAAGTATGGTTGTAATTTGAAAAATAAAAAATAGGAC
 AGAATTCATGTTGAATCGACTGTCCCACTCTTGTGAATGATATTACTTTTA
 GTGATTCTTTGACATAGCTATCTTATTAGACAATATTAAGTTTATTGACAA
 TTATCATCAAAAGCATCTTTTAAATCCATAGCTATATCATTGATGTCGCGC
 CCTTCAATATGTTTCGCGAGGTATAAAATGAACATAATGTTGACCTTTAAAT
 35 AATGCGTATGAGGGACTTGAAGGAACTTGTGGATGTAATCTCTCATGTGTT
 TGTGTTGCTTCTTTATCTTGACCAGCAAATACAGTTACTTTATGATCTGGT
 TTCACTTCAATTTTGTCTCTGCAACTGCAACAGCTGCTGGACGTGCTAATCCT
 GCAGCGCAACCGCATGTTGAGTTGATTACAACAAATGTTGTATCATCATT
 TCTATATTTTGCATATACTGATTGACGTCATCACTCGTTTCAAGACTTGTG
 40 AATCCATTGTCTGTTAATTCAGCTCTCATTTGTTGTGCAAGTTCTTTCATA
 TAAGCTTCGTATCCATTATCATATAAAATGCTCCTCTACTAATAATTTGTTT
 TAATTGTACTATATAGAAAATTAATTGCTTTTAGCGAACGTGTTTCTATT
 TGTGGTAATAATTCATTCCAGTTAGCGTTATCTTCTTTATCAATTGATATA
 AAATCCAAAACATAGAATATTGATTTAACACCTTCTATTTCAAATAATCGA
 45 TTAATAAATTCAGGTTGACCTTCTTGAGCGGCAGTATATGTTGTAGATGAG
 TTATCTTGACGTGGTTCACTCAACGAACTTTTCATTGTATTATGATTAGGT
 GTTTCTGAAATTCCTATAATCTCCATGTGACATCCTCCTATTGTAACCAT
 TACCAATTTGTTTAAATTTGTTTACCAGCGTACATTCATTTATACAAAAT
 GATGTGCTTGTGTTTACCCTGAGTTTCTTATTTGTGCTTTTTTAAATGGAC
 50 ATTGGTGAACAATAAGTATTCAATCAATTGGTCGATGTGTTGAATTTGCTTGT
 GTTCTGTATTAGTGAGTATAGCTTTCACCTTCCCTTAGCTATAGACTATTTA
 TATTTATACCCATTTTACATGTTTGTATTATAGCATTTATAAGTTGGGAAT
 TTGAAATAATCAATTTTAGTTTGTAGAAAACGTAACTTAGACATTGTTT
 TAAGTTATTATTATTTACTCTTCTTTTTTAGCATTGCAATTAAGATGAAT
 55 TTTAAGTAAATAGTATAGTTAATTTTGGCGGTTCTAAACACCCGCTTTA
 ACAAATTTAGGAGGAAAGTTATATGTATAATGAAATATTTGGTATTGCGT
 CATTTATTGTTACATTCGCTTTAATGGTACTGATGTATCGCTGTTTGGTA
 AACAAGGACTAATTGCTTGGGTAGCAATAGGAACGATTATCGCTAATATAC
 AGGTCATAAAACGGGTTTCATATTTTGGTATTACGGCTACACTTGGAAATG

TCATGTTTGCTTCTATATATTTAGCTACTGATATATTAATGACATCTATG
 GTCGTAAAGTTGCTAAAAGAGCGGTGTGGCTTGGTTTCTCTTCTACCTTAG
 TAATGATTATAGTCATGCAAAATGTCATTGCATTTTATTCCTGCCCCAGTAG
 ACAATGCCGCAAACTCATTTAAAAATGATTTTGTATTTAGTGCCTAGAATTG
 5 CTATAGGTTCCATTATTTGCTTATATCATAGGCCAACATATTGATGTATTTA
 TATTCAGTATGATTAAGATATTTAGCTCTGATAAGACCTTTTTTATTA
 GAGCATATGGTAGTACCATTTAAGTTCTATCATTGATACCGGTTTATTTG
 TTTCAATTGCTTTTATTTGGTACTATGCCTGGTACTGCTGTTTTTGAAATAT
 TTATTACCACTTACTTGTAAAACTAGTGTCAACTATTTTAAATGTACCAT
 10 TTGGATATATCGCTAAGTCACTATATCGAAAAGGAAAGATAGAACAACCTAG
 ATAATGGGTATTGATATTTATTGTATCAGAAATATAGATATACAAAAGCGA
 ATCAGAGTTCCATGTGCTGAAGTACTGCTCGAAATTCATTTGGTTTGTCCG
 TAGAATTTTCGAGCAAAATTCCTCAATGATAGGACCTACGACCAAGTTAAA
 TAAATCTTAAATATAAGGGCTAAAATTTAGACATCAATGCCCTATTTGTAGC
 15 AGAACCTGAAGTATATATGTTATCTTAGGTTCTGTTTGTTTTTGATTTAAT
 TTTTAGTGTTTTGTCTTTTGTAGTTTGATACAATAGAATAGATACATTTTAG
 TGAGTAGACAATTCTGAAGAGGAGTAAAAGTAATGGCTAAAATACATTTTG
 ATGCTGCCGACTAAAGGAAATCCCGGCCGAAGTGTGTCGATTTATTATTA
 AAGAAAATTCACAAAGATATACATTTACCCATGATTTAGGTGAAATGGATA
 20 ATCATAGTGCAGAATGGGCAGCAATGTTACACGCTTTGGAACATGCACGCG
 AATTAAGTATCTAACGCGTTACTTTTTACTGATTCAAAATTAATTGAAG
 ATAGTATGATGCAAGGTAAAGTTAAAAATGCTAAGTTTAAAGTTTATTTTG
 AAAACATAGAAATCTTAGAGCAAAGTTTGTATTTGATGTTTGTGAGATGGA
 TTCCACGAAAGCAAAATAAAGAAAGCAATCAACTTGCTCAACAAACACTAT
 25 ACAAACCTTACATCATAATAATCTAAAGTGTTAATATGATATGTTACGTATT
 TAATTGTTTGTTCACCTTACCAATAAGATTTCCGGCTTTGAAAAGCTTATT
 ATAATATGTTTGTACTGAATTATCGATGGTTAACAAAAACAAGCTTGGGA
 CATTAAGTCCCGGGCTTGTCTGTGTTAAAGATATATTTTACTTTTATAGAA
 CGCTTTTGTGTTTTTACAACCTTTTTTAGTTGATTGAGGCGTTTTCTCTTT
 30 TTAGAAGTAACTTCTCTGTTTTGTATCTTTGTTATCACTGTTTTGTGTTA
 CGTTTTAACGACGGGGATTGAATGAATATCTCCTTCACTCAATTGCTTTTCT
 TTAGATTTCTTATGTTTAGCAACTGGAATTGAATGTTTGTATCGTTTTGA
 ATTTGTTCTCTTTATCTATTTTACGCTTAGTTATTACAATTGGATCTATT
 TTATCAGTATCTGAAAGACGAATGTCTTTGTGATAATTTTTTATCGATTGT
 35 TTTTCTTCTCTTTTTTACGCTTTTTAGCCAAAACCTAACAGAACTAGAA
 AGTAAACCACTTAATGTAATAACACCTACAGTTTGGCCGAAATTATCAACT
 AATGTTTGAATAGG

40 Sequence 3437

step.1003b07.cons.ok

CTTTCAGCAGCAAAGTTAGCTTCACCTGAAATTCATGTCTTTTCATGTTGAA
 CTCTAAGCAACATAACAATATCTACTTTATCAATGACTTCATCAATTTCAA
 CATAAGGCGCTCTAATGTATTATCTACCCATTCTTTTGGACTTGAGAACA
 45 TTACGTTGGCACCTAATGATGTTAAACTATGATAATTACTTCTTGGGACAC
 GAGAATTTTTAATGTCCCCACATATTAGAATATTCAAACCTTCAAACGAAC
 CATATTCTTCATATATTGTCAATTATGTCTAATAAACTCTGAGTAGGATGTT
 GTCCACTTCCATCACCTGCATTAGCAATTGGAATATTTAATTGATCCAGTT
 CTTTCGTAATAAGAATTTTGGGAGTGACGTATGACAAGTAAATCAACACCTA
 50 TACTTTCAAGTGTTTTACATGTGTCATAAAGTGACTCACCTTTTTTACAG
 ATGATGACTTGTGTTTCAAAATTAATAAGTTTTAATCCTAATTTTTGTCTG
 CCATCTCAAAGCTACACTTTGTTCGCGTTGAATTTTCGAAGAATAAGTTTG
 ATACGTATTGACCGTTAAATTGAGGTAATGGTCGCTCACCAGATTGAATT
 GGCAAGCGATAGTAATTAAATCATAAAATTTCTGAATTAGATAAATGCTCCA
 55 TTGATAATAAGTGTTCCATAAAGCGCCTCCTATAGAAATAAAGCTTAATTT
 ATTTGGTTTTTATCTTTTGGCAAAATTAATTTAATATTATTCCTGAAAGT
 GCTGATAATGCCATTCTTCAATTTGTAAATTGATACCTATGCCTTTTAAA
 TTGATAAGAAGATTACCAATCCCGACAACAAGCACAACTGATGCTATAACC
 AAGTTGCGATTGCTTGGCAAACTACTTGACTTTCAACAAGCATTTCTTAAA

CCACTAGCTGCTATAATACCGAATAATAAAATTGAGACACCACCCATCACT
GGCGTAGGTATTGAAGATATTAAAGCAGTGAACCTACCAATAAATGCAAGA
ATGATAGCTATAACTGCCGCACCACCAATAACGTAAATACTATATATTTTG
GTGATCGCTAGTACACCTATATTTTACCATAAGTTGTACTAGGAGGACCT
5 CCTATCATACTTGCAAACATAGTTGAAACACCATCACCATGATTGATTTA
TCTAAACCTGGATTTTCAAAGAAATTGCGTCTACTATTTTATTAATTACC
ATTTGATGACCAATATGTTCACTTACCGTCACAAACACCACGGGTATCATC
ACGAGAAATGAGTCCTAAATGAAAAGATGGTGTGTAATCTTTAAATGGTAGA
TAAATATGAGGAAAATCTATCCATTTTCGCTTGTGCTATTGGAGCAAATTTA
10 ACTATGCCCATGAAAATGGATACAATATATCCCACTATAATACCTATAAGT
ACAGGTATTAGTGATAAAAATCCTTTGAAGAAATCCTTGGACGATGATGGTT
ACTGCTAATGTAATCAAAGCAACAATTAAGTAACTTAAGTTATACCCTTTC
ATTTACGACAGAAATTTTCGAACATGGCCATGTTTACTGCTGTAGGAGCTAAA
CTTAACCAATGACCATTATTACTGGTCCGACAACTACTGGTGGTAACAAG
15 TGCATTAACCAATGTGTTCCACTCAATTTAATGAAAATACCTATAATCAG
TACATTAATCCACTCATAAATAATGCAACCAGCATATCTCCAAGACTATGT
GTACTTAACCCTGTAATGATAGGTGTAATAAAGGCAAACTCGATCCTAGG
TAGGCAGGAATTTTCGCTTTTGTGATTAATATGTATAAAAGTGTACCAATA
CCTGATGCCAACAAAGCTGCTGATATAGGCAGATGTGTTAAAAATGGTACA
20 AGTACAGTTGATCCAAACATTGCAAATAAATGTTGAAGACTCAGAAATGCC
CATTGTGCTGGTTTTTGGTTTTATCTTTTACATCAAGAACTGGCTGAACGGTA
CGCTCAAACATTTGCTCATTTTCCATGTTGATTCATTTCTTTTCATAAAAA
AAGTCTCTTTACAATCAGTAGAATGTAAAGAGACTAAAATGAAATATATCT
AAGTGTAATGTTAAGTTAACTTTTATCATCACATTTATCAATGAAGCTTC
25 TATTGGATATTTATAGTTTACATAAATCTTTTTATAGTCATATGTAGATT
TCCATAAGAAACACTTCACCTTTTTTCAGTCTCTCGTACTGTTTTAAAGGA
TAATTATTCAATGACGACTGCATTTCTATCGTCAATTTCTTCTAAATAAAC
TGAAACAGATTCATCTCGTGTGTAGGTATATTTTTCCTACAAAATCTGC
GCGTATAGGGAGTTACGATGACCACGATCCACAAGTGCTGCAAGCCCTAT
30 TTTAATAGGTCCTTGTATGTAATAAAATCGCATCTAATGAGGCTCTTACTGT
ACGTCCGGTATACAAAACATCGTCAATGATAACAACTTGTGTTATTAAT
ATCTACATTAATATCAAAGCGTATTGATCAGCTTGTGCACTACCTTATC
AACGTCATCTCGAAATGCGTGATATCGATAGTACCTGTTGGTACTAATTG
TTGTTCAATTGAATTTATTTTATCTTGTATACGATGTGCTAAAAAAGCACC
35 TCTTGTTTTAATGCCTAATAGAACTAAATCTTTAGTTCCCTTGTATATTC
TAGAATTTTCATGAGCAATTCGTGTAATTGTACGTTGTATCGCTGCCTCATC
TAAAATAAATTCGTTAGACATTTGTTTACCTCTTTTAAATATAAAAAAATC
ATATCTTTAATAAAGATATGAGACAAAAAGTTAAAGAGAATAATCTATATT
CTCTATTATATTACATTTCTTAAGTTATAAACACAGCAAAAGTATTAGTAA
40 CTTAATCATCAAAATATTATTATCCTAATTATTCATATCTTCGGAGCCTCT
CTGGACCCCTCAATTAAAGATGACGTTTTTATATTAAACTGTTATTAGTTT
AAAACCTTTTAATAGTATAAGTCAATGGCAAAGTTAGTATTTATAAATTA
ATCGATGCCTCATTCGTTTATGCATCTCTACGCTTATATCATCTAGTAAC
TTTTCAAAATCGTTAGGTAATGTTGCATGCTTCTCAATGTATTTCATGTGTT
45 ACAGGATGTTGAAACCAATTATACCAGCATGTAAAGCTTGTCCATCGATA
TCTAACGTTTTTTTAGGTCCATACTTCGGATCCCCCTACAAGAGGATAACCA
ATATATTTTCATATGTACACGAATTTGATGTGTACGCCCTGTTTCTAATTGA
CATTCTATCAATGTATAATCTTTAAATGCTCTAATACATTAAAGTGTGTC
ACAGCTTCTTTACCGTCATCAACCACAGCCATCGATTGTCTATCGTTTTTA
50 TTTTCGACCAATTGGTGGTCAATAGTGCCATAATCATGAGGGATATTACCA
TGAATAACGCGGTATTTTCGTTTAAACAGTTTATAGACATAAGTTGTTCA
ACTAAATGACGATGAGCAACATCATTTTATAGCAACCATTAAATAAACCAGAC
GTATCTTTATCTATTCTATGAACTATGCCGGGACGTATTTCCGCCATTGATT
CCTGATAAATCTTTAATTTGATACATTAATCCATTTACTAATGTTCCACTA
55 TAATGGCCTGGAGATGGATGTACAACCATACCTTTTGGTTTATACCTTAGA
GCACAGTGCGC

Sequence 3438

step.1003b08.cons.ok

CTGTAATTATGAGTGAACGACTAGTTAAAGATGATGTATATACATCTATTC
 ATATTGAAGAGTATGAGTCAGAAGCTCGTGATACGAAATTAGGACCTGAGG
 AAATCACACGTGATATTCCAAACGTATCTGAAAGTGCACCTTAAGAACTTAG
 5 ATGATCGTGGTATCGTTTATGTTGGTGCAGAAGTTAAAGATGGTGACATCT
 TAGTAGGTAAAGTAACGCCATAAGGTGTAACCTGAATTAACAGCTGAAGAGA
 GATTATTACACGCTATCTTTGGTGAAAAAGCACGTGAAGTTCGTGATACAT
 CATTACGTGTACCTCATGGTGTCTGGCGGTATCGTTTAGATGTTAAAGTCT
 TTAATCGTGAAGAAGGAGACGACACTCTATCTCCTGGAGTTAACCAATTAG
 10 TACGTGTATATATTGTCCAAAAACGTAAAATACATGTTGGGGACAAAATGT
 GTGGTCTGTACGGTAACAAAGGTGTTATTTCTAAGATTGTTCCAGAAGAAG
 ATATGCCCTTATTTACCAGATGGTTCGACCAATCGATATTATGTTAAACCCAT
 TAGGTGTACCATCACGTATGAACATTGGACAAGTATTAGAATTACATTTAG
 GTATGGCAGCTAAAAACTTAGGAATTCACGTTGCTTCCCCAGTGTTCGATG
 15 GTGCCAACGATGATGACGTTTGGTCAACTATCGAAGAAGCTGGAATGGCGC
 GTGACGGTAAAAGACTACAAAGTTTAGGTCATAGATGTTAAAGTGATGGATG
 ATCGTATTTTCAGTTGGTGTCTATGTACATGCTTAAACTTGCACACATGGTAG
 ATGATAAGTTACATGCACGTTCTACTGGACCATACTCACTTGTACACAAC
 AACCATTAGGTGGTAAAGCACAATTCGGTGGTCAACGTTTCGGTGAGATGG
 20 AGGTATGGGCACTTGAAGCTTATGGTGTCTGCTTATACTTTACAAGAAATCT
 TAACTTATAAATCTGATGATACTGTAGGTCTGTGTTAAACTTACGAGTCTA
 TAGTTAAAGGTGAGAACAATTTCTAGACCTAGTGTCCAGAATCATTCCGAG
 TATTGATGAAAAGAACTACAAAGTTTAGGTCATAGATGTTAAAGTGATGGATG
 AGCATGATAATGAAATCGAAATGGCTGATGTTGATGACGAAGATGCAGCAG
 25 AACGCAAAGTAGTAAAATATACTGGATTAAATAACATAATTTGATGAATCA
 AACCACCTACGCCGTGATTTTGGGCACCCATAAAATTGGAGACATATAAAA
 ACAATATTTCTCATTAATGCTTGCATAATCATCTGCGTATCTCTAATTAATA
 TCCCCAAAGTGGATGTTAAAGTGCCACCGACGATGTTAGCAAATATGCAA
 AAGGTACATATAAGTAATTTGTACAATGTGAATGAAGGGATAATTCAT
 30 TGAACATACAAGCTATTATAATAATTGCTAATAATCCTAAATGACCATAGA
 ACCTACTTGTACAATATAAGTAGGAATGATTGAGAGTGGGAAATTCATCT
 TTGCCACTTGATTGAATTTCTGTGAGATTGATTAGTTCCTTCTAGGACAC
 CTTGATTAATAAAGAACCACATACTTATACCAACTAACAACCAAAATACAA
 ACGGAATGCCGTGTATAGGTGCATTGCTTCTAATTCCTAAACCAAAAACAA
 35 GCCAATACACCATGATTGTAAAGCAGGATTCAACACTTCCCATGCAATAC
 CTAGATAGTTATTATGATTGGTAATTTTAATTTGGAATTGTGCTAGTCTTT
 GAATTAAATAGAAGTTTAAATATGTTCTTTTAAACCGTTTCCAACTGCGA
 ACATTAAACTTAACCACGCTTCTCTCTATAATAGATTATCTATCTATTGTA
 AATCATTAATAGTTTCTCTTAAATTTGATTTCATTGAGTCCATTCTTAAATA
 40 ATAACATTTAACTGTTTATCATAAGCTCATTATATTATACAATCAATTCAG
 TTAAAAATTTGTTAGAAATTTTCGTGTAAGATTAAATTTAGAAAAATGAATCA
 CTTTCTTAACTCACTGTGTATATTTATTTTCAACAAAATCTTCTATTTCAC
 TTTACACTAGGTTAATAAATTTAACAACGTATGTGTTGTATATTATAATCT
 ATATAAATGTTTATAAGGAAGGTACAAAAATGAGCGTTTCGGTAAATATTG
 45 AAAATTTAACAAAAGAATACCGTATCTATAGAAATAATAAAGATAGAATTA
 AAGATGCATTAAATACCTAAAAACAAAATAAAACATTTTACGCTCTGGATA
 ACGTAAGTTTAAACAGCGCATGAGGGAGATGTGATAGGTTTAGTCGGCATCA
 ACGGTTCCAGGTAAGTCTACTTTAAGTAATATGATTGGTGGCTCTATTTTAC
 CAAGTTCCGGTGAAATAACGAGACATGGTGATGTGAGTGTGATCGCTATTA
 50 ATGCAGGACTAAATGGACAATTGACAGGTGTAGAAAAATATTGAATTTAAAA
 TGCTCTGCATGGGCTTTTAAAGGAAAGAAATTTAAAAATTAATGCCGGA
 TTATAGAATTTAGTGAACCTCGGCGAATTTATTTATCAACCTGTAAAAAAT
 ATTCAAGTGGTATGCGTGCAAACTTGGATTTTCAATTAATATTACTGTTA
 ATCTGACATATTAGTTATTGACGAAGCATTATCAGTAGGCGATCAAAACAT
 55 TTAATCAAAAATGTTTATAGATAAAATTTATGAATTTAAAGCGGCTAAAAAA
 CAATATTTTGTGTAGTCATAATATTAGACAAGTGGTGAATTTGTACAA
 AAATCGCTTGGATTGAGGCGGTAAACTAAAAGAATTCGGCGAACTTGAAG
 AAGTATTACCTGATTATGAGGCGTTTCTTAAACTTTTAAAGAAAAATCTA
 AAGCAGAACAAAAGGAATTTAGAAATAAATTAGATGAGTCACGTTTGTCTG

TAAAAATAAATTTGAAGAAAGTAAACCTTAAACATTAAGGGAGTGCAATAAA
ATTCATCCGAATTGATTGCCTCCCTTTTTTACTCTTATTTTGATTTCATTT
TATATTTTTGTATTGCTACCTTTAACATAAAATTTAGGAATACTTATCATTC
TCCCAATGCGTTTCCAATCAATAAGCACACGATATACCCACTCAATATTTA
5 ACTTTCTAAATATTTGAGGTGCTCTTTTCTTTGAGCCACTGAATACTTCAA
ACGACCCACCTACGCCCATCATCACAGTGTGCTTAACTTGTCCTTATGCT
TTTGAATCCATTGTCTTGCTTTGGAAATCCCATTCCCTACAAAAATGTAAT
CGGGATTAAAACTTGTATACGTTTATGACTGTTTCATCTTCTAGATGAA
TATAGCCATGATGATGTGCAAAGTGGATATTAGGGTATTGAGAT

10

Sequence 3439

step.1003b09.cons.ok

15

AGCAGAAGTAGCTGAAAACAACATTGTGAAGGAAAACCCAGATGTAGAAAT
CTTTGAAGAAGGTATTATTGATTTCATTTCAAACCTGTAGGTTACTTTTAGA
AATTCAAAATAAATTGGATATTGAAGTGTCAATTATGGACTTTGATAGAGA
TGAGTGGGCAACTCCTAATAAAATTGTAGAAGCTTTAGAAGAGTTACGATG
AAACTGAAACCATCTTACCCATTATAATCAGTGGCGTCATATTTATTATA
TTTTTATTTTTTACCAGCAAGTTGGTTTACTGGTCTTGTGACAAAATAAGACG
20 TTAGCTGATAATAGAATTTCTTTAACTGATCAAGTTTAAAAGGCACGCTC
ATTCAAAATAAATTATTCGAGTCAAATAAGTATTACCCGATTTTATGGTTCA
AGTGAACCTGGAAAAGGATGATCCTTTTAATCCAGGTATAGCGCTAAATAAA
CAAAATGCTAGTAAAGAAACGTTCCCTTATTGGTGCAGGAGGTTGCACAGAT
TTAATTAATGCAGTTGAACCTGCAGCGCAATACGACAATTAAAAGGAAAG
25 AAATTTACATTTATTTATTTCTCCACAGTGGTTTACGAATCATGGATTAACA
AATCAAACTTTGATGCTCGTATGTCACAAAGTCAAATCAATCAGATGTTT
AATCAAAAAGATATGCCAGCTAATTGAAGAAACGGTATGCACAAAGATTG
TTACAGTTTCCGCATGCACACAATAAGTCATACCTTAGAGAACAAGCAAAA
CATCCTAATGATGTCTCTGGAACTACATTTCTTCATTTAAAGAAAATCAA
30 TTAAC TAAGATTGAAGCTATTAAATCATTATTCTCATTCCTAAGCCACCT
CTAGCAGAAGTAAAACCTGCAACAAGAGAAGATGCTTCATGGGATGAGATG
AAACATAAAGCTGCCGATATAGGCAAAGCAAATACTCAATCTAATAAATAT
GATATAAGAGATCCATATTGGAAATTGATAAAACAAAACAAGCGTAAATC
AAAAGGGATTATGAGTTCAACATTAAC TCACCCGAGTTCCAAGATTTAAAA
35 TTATTAGTGCAAACGCTACATGCTGCTGGAGCTGATGTACAATATGTTTGT
ATACCTTCAAATGGAAGATGGTATGATCATATAGGTATCAAAAAAGATAGA
CGTGAAGCTGTATATAAAAGATTCACTCAACTGTAGTTGATAATGGTGGG
AAAATTTATGATTTGACAAATAAGGACTATGAAAAGTACGTAATTAGTGAT
GCTGTTTATATTGGATGGAAGGTTGGGTTTACGTCGACCAGCAAATTGCA
40 AGACATATGGATGGTTCATGCGCCTAAAAATCATGAAGTCGATTATTTCTAAA
AATAAACCCCGCAACAAACATCACAACGATCGTCAAGATGATCAACATCAA
GGCAACAAAATAATATTATTAACCGAAAATAAAATTAAAGTGAGCTAGGTACAG
AATTGCTTATTTTCATCCCGTGTAAACGAGATTTATAAAGGAAATACCTCGAC
GTAATGTATTTTTCGTTTACTTCTTTGATAGCCAGTATATAAAATGCTTAG
45 GAACAATGTTTGTATCCTAAGCATTTTATTGTTTATTTAGCATCATAGA
GTTAATATTAGTGCGCTATGACTTCATTTATTAATTCTATCAAATTA AAA
AACCTAGGACGTCATCTATCTTACCTAGGTTAATAAAATAATTATATAGGGAT
GTATTAGAATACTTGTCTACTTCTATTACACCTGGAACCTCTTCATGAAG
TGCACGCTCGATGCCAGCTTTTAAAGGTAATTGTAGAACTAGGACATGTTCC
50 GCAAGCGCCATGTAATTGTAGCTTAACAATACCATCTTCAACATCTACAAG
CGTACAATCGCCACCATCTCGTAATAAGAATGGACGCAAACGTTCTATAAC
TTCTGCTACTTGATCAAACATTGTTGGATTCTCAGTAGGCATGTGACATGT
CTCCTTTTCAGAGCAATTATTTTCAATAATCATTATTATTACATTATAATAGAT
AAAGTAGTAAAAATCTATAAAACAAGATAGGGGATTTTAAATGACTAAAATT
55 AGTGTGTGCTATATGGAGCAGAAGTCGTTTGTGCGAGTTGTGTAAATGCA
CCTCATCTATAGATACTTATCAATGGCTTCAAGCATTACTTTTAAAGAAAG
TTTCTCAACATCATTTTGAATTTACATATATTGACATACGAAATGATACT
GAAAATTTAACTGATCATGATATGCAATTTATAGAAAGAATTAATGAAGAT
GAATTGTTTTTACCCATTAGTTACGATGAATGATGAATATGTAGCAGATGGT

TACATACAATATAAAACAAATAACCCGTTTTATTAAATCATATTTTACTATG
 TAAGATTAAAAGAGGGGCAATATTCATACTAACGATGTGCCTCAATTGAC
 AGTGTCTGTGTATTTTTTAGTGAAAATCTTCATTTTTAGGCTATATCTTT
 TATTGAGAAGAAAGACATTTCAATGCTGTGCATAGTTTGTGATTATTAATTA
 5 AGCACCTAGGGAAATGTAAATGCCCTAGGTGCTTTATTTATTAACCATTAT
 GATATTTATAGCGCCACAGAACACCTGATTTTAGAATAGAGGCTAATCGTC
 CGGTAACGTGTTCCGATCCATGATATAAGCAAAACCTTGTTTGTACCTAAAG
 AGCCTAAAAGCCCTGTACTTTAATTTTCAGGCATTTTATCTGGAAGTGGTT
 CGTTATTCCATTGCTTCTTCAACACCTCAGCAATCTGTTACCTTGTAGTT
 10 CTGCTAGTTGAGCACTGGGTGCATGTGGTAAATTAGCACAGTCACCTACGA
 CATAAACATTTCTATAGGTTGGGACTTGATGGTACTGATTAATAATTACGC
 GTCCAGTGGTACTCATATCAATAGGAAGATTACGCACAATTTCAACAGGTT
 GTATGCCTGCTGTCCAAACGACTAAATCAATATTTTCTGGTTTACCATTAT
 TATAAATTTTCCGGGTTCTACTCTGTGCGATGACTGAATTAGGTACTACAG
 15 TAACATTTGTGTTTAGAAAACCAATTAGATATGTATTTACTCAGTTTCTCTG
 GAAAATTCCTTAAAATTCGAGGCCCTCTATCATATAACAAAATTTCCAAGT
 CTGATCGACTTTTACGTAGCTCGCTCGCTAATTCAATGCCACTTAAACCTG
 CCCCACGATACCTACACGTGCCCTTTAGGTAACCTCGCTAATTCCTATGGT
 ATGTTTCACGCGATTTAGATAATGTTTGAATGCTATGTGTATATGCTTCAG
 20 CACCAGGGACATTTATGATATTTATCCTCACACCTTAGACCAATGATAAGTT
 CGTCATAATCTATTTTGAATTTCCAACGTGTTATCATTTGTTTCGTCCAAAT
 CTATATCACTGATTTCCCATAAACCGTATTAATTTGACTGCTATCTGGAA
 ATTTGGATTTCGCACCTCTTTGTCAGATTTAGTTCTGCTGCAAGTGCATAAA
 ATTCAGGTTTTAAACCGTGAATGGCATGCGGTGATTAAGTTAAGTGAT
 25 ATCCCTCAGGAATTGAATGAGGTAAATGCGCGACATAATTTCGCATATTAC
 CATAGCCCCCGCTAGTAATACTAAGTTTTTCATTGATTCCTTAGAGCACA
 GTGGCGATGATATCAGAT

30 Sequence 3440

step.1003b10.cons.ok

ATTC AACCACTTTACGAATATATATAATAATAACGTTTGCATTACCATATT
 AACTATCTGAGTGATAGGAAATCTTAAAAACTTTTCTATTGTAGGTTTAC
 TTTATATACAAAATAACAATTTAAATAATATGAAATAATAAACTTACAAT
 35 AAATCCAACAATGTGACTAACCATATAATTACATGTAACACCTTTAACAA
 AAATAAGTATGTTATATAGTAGTTAAAGGTATTAATCCACCAACTATGAT
 AAACCTTTATAATCTCATAATGTATTCGGGTAACTTCATAATTTCCCT
 TTGAACAGTTATTGTTTGTGACCATATGGGACTACTTTAAATGTGATGTCT
 TTGGGCAAATCATAATAGTGTGTTGAAATCTTGAAGTTGTTCTTCGACTTT
 40 GGTGTAGTCATATGAGTATATCGCTCAAATGGTAATCTTTCCAATGTTATT
 TTCTGATAGCGATGATGTTGTGTTATTTGTTCTTTAATGAAGTCTATTCTG
 TGCACACTACTAATATGTATAAAAGTAAATCCAATCATCATAATGATGCTG
 ATGATAATCGCAAAATATTTTAAATATATGTTCAAGTTGTTTAAATAGCACA
 TCACATTGCTGAATTAAACAAAGTAATATCACGATCCATAAAGTATAAATA
 45 AAATAAAAAATTTCTATAACTTATAGGCGTCACAAATAAAAGTGGCAAAACA
 GATGAAGCCATAGCTATAAACTCCCATCACAATCATTCTTATGTATCTT
 TGCTGTATCATTTTAAACACAACGTATATCACACTTATCATGTAATGAAG
 CAAATCGTTGTTTCAAAACGGCTATAGAAAATGAAGCTTTATATAATTCA
 AAATGAAATTGATTGTAAACGAAGATCTTATAAATAGGTAAAGTAATTAAA
 50 CCTAAGAGTAGTGGTATTTTAAATATAAACTCTCATATGCTTCAGGCTTTTA
 TTTTGCCTTAAGTAAACTTACTTACTATTGATATCACGGTAAGAAATATC
 ATTTGATTAAATAACATATATTCTGGTACAAGCTTAAATAACGTACACCT
 GCTTTATGTATCATTTCCATGACTATCGGAAATTGAATAATGCGTATTTAAT
 CCATCCTTAATTAATAAAATAATTGAAGTTTAAAAACATTATAATGTTACCT
 55 ATACAACCTAAGCAATAAATCTACAATTAAGAAATAACTGAGTCTTTTTTTA
 ACAAGAAAATGACTACCATTCCTATTAATAATAATTAAAGCTATTAGCGATG
 GAAAGATTCTCCAAGAAGAATTGTCCAAACAACTTACTAATAAAAAATACC
 CATAATTGCATTTTCAGAACTGTATCGTGCGACTCAATCTTTTTAACTACC
 GTAAAAAGAATAAAAAAGTGATAGGACTGTAGCAGGTATATAACTAAAAAT

CCAGTAAACCACCCGTAAGTTTCGCTATAAAATTGTATTAGGTACAGTAACC
 ATTAACACAAAACCTCAAAATAAAATAAAATACGATTCTGTATGTAATTGAACC
 ATATAAGCAACTAAATAGATAACTAAAAACGAAGTTATCGCATATATTAAA
 GCTCTTATTATGTTATTATGTACGGCAACCCATTCAAATAAATGTCCATAA
 5 TAACGACCATTTCCTTGAGTCAAATATTGACTTAGATTAACTTTATATGCA
 TGCCAATCCGTAGTTGATAATGGTGTAGAAATTGCCATTATAAGATAAAAG
 ATAAAAATTCCTATTAAACCAAAAAATGATTCATATTGTTTAAAGTTATTC
 ATTGTTGACCTCTCTGTTCTTCATGATGTAATTCAAAACCTAACATTTAAAC
 GATTATTCATGTTTATTCTCTTTTATGTGATTATTCAATTATCTACTGA
 10 AAATATATAGACATTTATTGAAGCTTTAATCATATTAAATATATACATATTG
 CTCATCGTCAATTTTATATTCTTATTATTATACAATTCACGTATGATTTC
 ATTTTAAATTATCAAGATTTCTTTAGTTTAAATTTCAATTTTGGATAGAATAA
 ATATAAACCAATTTGTTTAGGAAGTGATATAAATGTCACCATATAAAGTTAT
 CATTGCACCTGATTCTTTTAAAGAAAGTATGTCGGCGAAGGAAGCTGCTTT
 15 AGCTATTAAAGATGGATTCCAAGAGGTGTTTCGATTCCAGTACAATATATGA
 CATTATTCCTATGTTGAAATAAGTCATCGTATTTCATGAACGCCCTGAATTA
 GGCAATGAAGAAATTTTGCATCGAGAACATTAATTGACCAATTAAGAGCA
 AATCGATTTCGAAATCGAAACGGATATTGCAGGACATGCAACAGGATTTATA
 GCAACGTATGATCTGTATGATGACTGGACCGGTTATAGGATTTCTAGCTGAA
 20 TATGATGCTTTACCTGGTCTTGGTCACGCATGCGGGCATAATATTATGGT
 ACTGCTAGCGTACTTGTGCTGAGTAGCACTAAAAGAAGTCGTCGATGAAAT
 GGTGGTAAAGTAGTCGTTTGGGATGTCCTGCTGAAGAAGGTGGGGAAAT
 GGCTCCGCAAAAGCTTCTTATGTTAAAGCAGGTGTCATTGATGAAATGAT
 GTAGCATTGATGATTTCATCTGGAAATGAAACTTATCGTACAATTAATACT
 25 TTAGCTGTGGATGTTCTAGATATTAAATTCATGGACGTAGTGCAGCATGCA
 TCTGAAAATGCGAGATGAAGCATTAACGCTTTAGATGCAATGATTTCATAT
 ATTAATGGTATAGCACAGTTAAGGCAACACATTAAAAAAGGACAACGAGTT
 CACGGGGTTATTTTAGACGGTGGTAAAGCGGCTAATATTATACCTGATTTT

30

Sequence 3441

step.1003c04.cons.ok

GTATAAGTAGCTAATTTTTCATCGTATAAAGTGTAAGGAGATTGTCTGCCA
 35 TTGACAATAGCGTTCCTTTAAATAATTTAATTCTCACATCTCCCTCAACA
 TATTGTTGAGTACTATCGATAAAGAGTTTAAACTATCTGTTAATGGCGAG
 AACCACAAACCATTGTATATTGTTCTGAAAATGTTTTCATGACAGGC
 TTAAAGTGCCTACGTCTTTAGTTAATGTAATTGTTTCTAGTGCTTTGTGT
 GCTTTTAAATAAATTCCGCACCAGGTGTTTCATAAATCTCTCTCGATTTT
 40 ATCCCAGCAATTTCTGTTTTCAACATGATCGATTCTACCAATACCGTGTTG
 CCAGCAAGTTGATTCAAGTAAAGAAATAAGGTCATCTAATTGATAATCTTTG
 CCATCAACTTGACTGGAATACCTTGTTTAAATGTAAGGATAAATTCGTCT
 GCATTGTCTGGAGTTTCTTCTAAAGGTGTAGTTAAATCAAATGCATCTTCC
 GGAGGTGCGGCATACGGATCTTCTAAAATACCACATTCATTAGCTCTCCCC
 45 CATAAGTTTGGTCAATTGAGTATGGCGAGTCATAATTGATTGAAACAGGA
 ATATTATGTTTGATTGCGTAATCAATTTCTTCTCTCTGCTCCAAGCCCAT
 TCACGAACAGGTGCAAATGCTTTTAACTTAGGATTTAAAGCTTTGATTGCC
 ACTTCGAAACGTAATTGATCATTAACCTTTACCAGTACATCCATGCGCAATA
 CCAATAGAAATTTGTTTTTTCAGCAATTTCAACCAGTTTTTCGCGATGAGT
 50 GGACGTGATAATGCTGAAACTAGAGGATATGCATTTTCATACATTAAATTT
 CCTTTAATAGCATAAATTACATAATCATCACTAAATTTCTTTAGTTGCATCA
 ATAATATGACATTGACTGCACCCATATCTAAAGCTTTTGTATATACAACG
 TCTAAATCTTTGCCTTCGCCTACGTCAAGACAACAAGCAACTACATCATAT
 CCTTTATCAATAAGCCATTGAACTGCAACGCTTGATCTAAACCACCTGAA
 55 TATGCTAAACGATTTTATCTTTTCATATTTACACCTCAAATAATTTTAT
 GATTATTAAATGCATCTATCATATCAGATTGATAATGAAAATAAAACCAT
 TATTTTACATAAAATATTCATTTTAAATCTGATTATATTTAGATATTTTGT
 TTTGAAATACTGTTATAAAAGCTTTTATATATATTTTAAATTTCTGTAAA
 TTGTATATCTAAAGTGGTGGTGTGACCTTTGATATTAAAAATAACTTTT

ATGAATGAAAGGTTACATAAAAAATAATTTTCCTGTGATAAGTTAGCCTTG
 AGAATTGAATAGCGCATCTAAACCTTAGTAGAATTAGTATAAGTACTTATAA
 CAGGAGGCAATAAGATGACTCACATTCAATTAGACTATGGCAAACTTTAG
 AATTTTTTGATAAGCATGAACTAGATCAGCAAAAGGATATTGTTAAACTA
 5 TCCATCAAACCTATTCTATAAAGGTACAGGAGCAGGTAATGACTTTTTAGGTT
 GGTAGATTTTACCTGTTGATTATGATAAAGAAGAATTTCTAGAATCGTCG
 AAGCATCTAAACGTATCAAATCAAATTCGGATGTACTTGTGTTATCGGTA
 TTGGAGGTTCTACTTAGGTGCACGTGCTGCAATCGAGATGCTTACATCTT
 CATTTAGAACAAAATACGGAATACCCGTGAAATTTGATTTGTAGGTAATCATT
 10 TATCCTCAAGTTATACAAAAGAATTACTTGATTATTTACAAGGAAAAGATT
 TTTTCAGTTAACGTTATTTCAAATCAGGTACTACGACAGAACCCAGCAGTTG
 CATTTAGATTATTTAAACAATTGGTTGAAGAAAAATATGGAAAAGATGAAG
 CTAAGAAACGTATTTTGAACGACAGATAAATCTAAAGGTGCACCTAAAC
 AATTAGCAGACAATGAGGGTTATGAGACGTTTGTGTACCTGATGATGTGG
 15 GAGGTTCGTTATCTGTTCTTACAGCTGTAGGATTACTACCAATTGCAACTG
 CAGGTATCAATATTGAATCAATCATGATTGGTGCAGGCTAAAGGCAGTGAAG
 AGTTATCTTCTGATGATTAGATCAAAATATCGCATATCAATATGCAACTA
 TTCGAAATATTTTATACAGCAAAGGTTATACTACTGAAATGTTAATTAATT
 ACGAACCCCTCTATGCAGTATTTCAACGAATGGTGGAACAATTATACGGTG
 20 AATCAGAAGGGAAGATTTCAAAGGTATTTATCCATCAAGTGCGAATTACA
 CAACTGATTTACATTCCTTAGGACAATATGTTCAAGAGGGCCGTCGTTTCT
 TATTCGAGACAGTGGTTAAGGTCAACCATCCAAAACATGATATCAAAATTG
 AAGAGGATGCAGATGATTTAGACGGACTGAACTATCTTGCTGGCAAAATCAA
 TCGATGAAGTGAATACTAAAGCATTTGAAGGTACATTACTTGCACATACCG
 25 ATGGTGGCGTTCCAAATATCGTTGTAAATATTCCTCAGTTAGATGAAGAAA
 CATTTGGATATGTTGTTTATTTCTTTGAATTAGCTTGTGCAATGAGTGGAT
 ATCAATTAGGTGTTAATCCATTTAATCAACCTGGAGTTGAAGCCTATAAAC
 AAAATATGTTTTCGCTATTAGGTAAACCAGGCTTTGAAGATAAGAAAAAAG
 AATTAGAAAATCGTTTATAAATTTTAAGTTGAACCTAAAAATGAGCCTAGGA
 30 TTATCTAGGCTCATTTTATTATGTGTATTAATAAGAGTAAATGAAAATTG
 ATTAACGTCAGATTAATAAGAAGTCAAATTCAGATAGGAAATATTAAATTT
 TCAATTCATTAGAGTTTCGATAAAGAAATGGAATAATAAAATTGATGAATTA
 GATATCAATATATAGATTTATTAATTCCTACCTAAGTCTGATTTATGTTGT
 TGACCTAATTTTAAATGTTTCAAAAAATATCATTTACAACCTATAATGAATAA
 35 GTAATACAACCCTAAGAAAGGATTTGATGTATTGTCGCTTTCTCAGTTAGA
 AGAATGGTTTGACGCATTTTCGACAATTCGGATATATTCCTGG

Sequence 3442

40 step.1003c05.cons.ok
 ACATTCAAAACCTAGATAGTAAGATTTTTCGTCGCAAAACGTTTTTTA
 AAAATTGATTAAAGTCTTCGATCGATTAGTATTTCGTCAGCTCCACGTGTCAC
 CACGCTTCCACCTCGAACCTATTAACTTCGTCATCTTCGAGGGATCTTATA
 ACCGAAGTTGGGAAATCTCATCTTGAGGGGGGCTTCATGCTTAGATGCTTT
 45 CAGCACTTATCCCGTCCATACATAGCTACCCAGCTATGCCGTTGGCAGGAC
 AACTGGTACACCAGAGGTATGTCCATCCCGGTCTCTCGTACTAAGGACAG
 CTCTCTCTCAAATTTCTACGCCCACGACGGATAGGGACCGAAGTGTCTCAC
 GACGTTCTGAACCCAGCTCGCGTACCGCTTTAATGGGCGAACAGCCCAACC
 CTGGGACCGACTACAGCCCCAGGATGCGATGAGCCGACATCGAGGTGCCA
 50 AACCTCCCCGTGATGTGAACCTTGAGGGGAGATAAGCCTGTTATCCCCGG
 GGTAGCTTTTATCCGTTGAGCGATGGCCCTTCCATGCGGAACACCGGATC
 ACTAAGTCCGCTTTTCGACCCGTGCTCGACTTGTAGGTCTCGCAGTCAAGCT
 CCGTTATGCCCTTACACTCTATGAATGATTTCCAACCATTTCTGAGGGAACCC
 TTTGAGCGCCTCCGTTACCTTTTAGGAGGCGACCGCCCCAGTCAAACCTGCC
 55 CGCCTGACACTGTCTCCCACCACGATAAGTGGTGGGGTTAGAAAGCCAAC
 ACAGTAGGGTAGTATCCCACCAACGCCTCCACGTAAGCTAGCGCTCACGT
 TTCAAAGGCTCCTACCTATCCTGTACAAGCTGTGCCGAATTTCAATATCAG
 GCTACAGTAAAGCTCCACGGGGTCTTTCCGTCCTGTGCGGGTAACCTGCA
 TCTTCACAGGTACTATGATTTACCGAGTCTCTCGTTGAGACAGTGCCCAA

ATCGTTACGCCCTTTCGTGCGGGTCGGAACCTACCCGACAAGGAATTTTCGCT
ACCTTAGGACCGTTATAGTTACGGCCGCGCTTTACTGGGGCTTTGATTTCGT
AGCTTCGCAGAAAGCTAACCCTCCTCTTAACCTTCCAGCACCGGGCAGGCG
TCAGCCCCATACATCACCTTACGGTTTAGCAGAGACCTGTGTTTTTGATA
5 AACAGTCGCTTGGGCCCTATTCAGTGCAGGCTCTTCTGGGCGTGAACCCATAA
GAGCACCCCTTCTCCCGAAGTTACGGGGTCATTTTGCCGAGTTCCCTAACG
AGAGTTTCGCTCGCTCACCTTAGAATTTCTCATCTTGACTACCTGTGTCGGTT
TGCGGTACGGGCACCTGTTATCTATCTAGAGGCTTTTCTCGGCAGTGTGAA
ATCAACGACTCGAGGAAACAATTTCTCTCCCCATCACAGCTCAGCCTTAT
10 GAGTGCCGATTTGCTTAACACTCAGCCTTACTGCTTGGACGTGCACCTCCA
ACAGCACGCTTCGCCTATCTTACTGCGTCCCCCATCGATTAAAACGATAC
TAGGTGGTACAGGAATATCAACCTGTTATCCATCGCCTACGCCTGTGCGCC
TCAGCTTAGGACCCGACTAACCCAGAGCGGACGAGCCTTCTCTGGAACC
TTAGTCAATCGGTGGACGGGATTTCTCACCCGCTTTTCGCTACTCACACCGG
15 CATTCTCACTTCTAAGCGCTCCACATGTCTTGGCATCATGCTTCGACACC
CTTAGAACGCTCTCCTACCAATTGTCCAAAGGACAATCCACAGCTTCGGTAA
TATGTTTTAGCCCCGGTACATTTTCGGCGCAGTGTCACTCGACTAGTGAGCT
ATTACGCACTCTTTAAATGATGGCTGCTTCTAAGCCAACATCCTAGTTGTC
TGGGCAACGCCACATCCTTTTCCACTTAACATATATTTTGGGACCTTAGCT
20 GGTGGTCTGGGCTGTTTTCCCTTTTCGAACACGGACCTTATCACCCATGTTCT
GACTCCCAAGTTAAATTAATTGGCATTCGGAGTTTGTCTGAATTCGGTAAC
CCGAGAGGGGCCCCCTCGTCCAAACAGTGCTCTACCTCCAATAATCATCACT
TGAGGCTAGCCCTAAAGCTATTTTCGGAGAGAACCAGCTATCTCCAAGTTCTG
ATTGGAATTTCTCCGCTACCCCTCAGTTTCTATCCGCTCACTTTTCAACGTAAG
25 TCGGTTTCGGTTCCTCCATTCAGTGTACCTGAACCTCAACCTGACCAAGGGT
AGATCACCTGGTTTTTCGGGCTACGACCAAATACTCAACGCCCTATTCAGAC
TCGCTTTTCGCTGCGGCTCCACATTTGCTGCTTAACCTTTCATCAGATCGTA
ACTCGCCGGTTTCTTACAAAAGGCACGCCATCACCCATTAACGGGCTCT
GACTACTTGTAAGCACACGGTTTCAAGTTCTCTTTCACTCCCCCTCCGGGG
30 TACTTTTACCTTTCCCTCACGGTACTGGTTCACTATCGGTCACCTAGAGAG
TATTTAGCCTTAGGAGATGGTCTCCAGATTCCGACGGAATTTACGCTGC
TCCGTCGTACTCAGGATCCACTCAAGAGAGAATATGTTTTCGACTACAGGA
TTATTACCTTCTTTGATTCTATTTCCAGATGATTCGTCTAACATGTTCTT
TTGTAACCTCCGTATAGAGTGCTTACAACCCCAACAAGCAAGCTTGTGGT
35 TTGGGCTCTTCCCGTTTCGCTCGCCGCTACTCAGGGAATCGATTTTCTTT
CTGTTCTCCGGGTACTAAGATGTTTCAGTTCTCCGGGCTGCTGCTTCTGAC
ATGCTATAAATTCACATGTCGATAACATGACATAACTCATGCTGGGTCCC
CCATTCGGAAATCTCTGGATCAACGCTTACTTACAGCTACCCAAAGCATAT
CGTCGTTAGTAACGTCTTTCATCGGCTTCTAGTGCAAGGCATCCACCGTG
40 CGCCCTTAATAACTTAATCTATGTTTCCACCATATTTTGAATGTTATTCA
AAATAAATAGCTAAAACCTAGTTATTAATCTTGAGAGTGTCTTTTCGAACAC
TAGCGATTATTTATGAATTCAGCTTATTTAAACTCTATTCACTCGGTTT
TGCTTGTAATAATCTTACTTGTCTATCTAGTTTCAATGTACAAATGAATG
TTAATAAACATTCAAACTGAATACAATATGTCACGTTATTCCCTCATCTT
45 CGTAGAAGATGTTCCGAATATATCCTTAGAAAGGAGGTGATCCAGCCGCAC
CTTCCGATACGGCTACCTTGTTCAGACTTACCCCAATCATTTGTCCCACC
TTTCAGCGCTAGCTCCAAATGGTTTACTCCACCGGCTTCGGGTGTTACAAAC
TCTCGTGGTGTGACGGGCGGTGTGTACAAGACCCGGGAACGTATTACCGT
AGCATGCTGATCTACGATTACTAGCGATTTCACTAATTTCTTGAAAAGTTT
50 TTTCTAATTTTTCAAAGAAACACCTTCTTGATGTGCTCGTTTGAGGTTT
TTTGATTGGGTCTTTGATTTTGTCTGTCTCGTGACTTTTATCTCGTAATG
CGCTAACACCTGTAATTATCACAGATATGATGAAAATGATAATGCCAATAT
TCATCTTATCACCTCTTCAATTATGTTGTGGTGATTCATCATCGTTTGTGAT
CAGTGCGCTTATTTATTGCAATTTCTCATACCTGTATCCGCTTCA
55

Sequence 3443

step.1003c07.cons.ok

TG TAGCAATACCACCATGCAACTGACCTACGTTTTCAAATCCTTCTTTTAC

963

ACATAACATTAAATCGTAATTTTACGATTTAGGAACGATATAGAGATGAT
 GGTGGATGAATATAAGATGGACATATCCGTCCCTGTCTATTTTATTTGAA
 GTGCAAAACAAAGTGAAAATTCTTCAATACCTCATATAATACATCAATTGAC
 CATCATCTATTATATTGTTCTTTTATGTTTAGTTTCCTTTTATTTCTAAC
 5 TCTTTTTTTTTTCACTCGTCTTAGGACGGGTGTTTTTATAGTAGAGTTAT
 GAAATTGTTGAAGGTTCCGTTTCGATTATATTCTGAAATTAAATCAATTCAT
 AGACTTCTTCGTATCTATGATGACTTGCAACATGGAATTTCAAATACCCCT
 TAGTACAATGAACATATTTATATAAAGAGTGAATGGGTAAACTACGTTT
 TAAAAATGTATAATGAGATGATATTTGAGTTTGTTTTATATAAAGAGAGG
 10 CAAGTTTGTTTTTTATATAGAAACGGTGTATATAAAAAATAAATTGTGACG
 GAATCGTGTGATTAGTTATTTTAAATTATTGAGTTTAAAAAAGGACAA

Sequence 3444

15 step.1003c08.cons.ok
 AAATATAGATTGTCGCTTTAACAACAGAATTCAAATCAGAACCTGCTTCTT
 TTAATACCACAGTTAAATTTTCTAAAACTTGCCTTAGTTGTTCTTGAACAT
 CATCGCTAACAAATTGTTCCATCAAGTGTGAGTGGAAATTTGACCTGATGTAA
 AGACAAAACCGTTTATAACAGTTGCATGCGAATATGGGCCTAGTGCCTCGG
 20 GTACCTTATCTGAGTTACAACAACCTTATTCCTAAATATTGTGTTATTTTCC
 CTTAAATCTCCTCAAAGATAAATGATTTAAGGGTTTTTTTATGCATAAAATC
 ATTTGTTCTAACCCGTTTTACTATCTATAAGATTGTTATTGTAATAAAATA
 TTGACATTAAATACATAACGAAAAATAATTGAAAGGAATGAAATAAT
 GGTACAAAATGCATTTGTTGCACTGGACTTTGAACTGCTAACAGTAAACG
 25 AACAAGTATTTGCTCTGTAGGTATGGTTAAAGTTATTGATAATCAAATAAC
 AGAGTCATTTTCATACCCTTGTGAATCCTTTTGACTACTTTACTGAAACAAA
 TATTACTGTACATGGCATCCATCCCGAAGATGTGCAAGATGCTCCTGGATT
 CAAGCATGTCTATCCATATATGCTAAAAATTTATTGATCAACTTCTGTTGT
 TGCCCATAAATGCTGCTTTTGTATGAATGTGTTACATCAAAGTTTAAAAAG
 30 TGACATTAAATACACCTTCTTTGACATACTTTTGTTCATATCAATTAGC
 TAAACGAACAATTAATGCATATCGATATGGGCTAAAAACATTTAATGAATCA
 TTACCATTTAGATTTTACGGTCATCATGATGCGTTAAATGATGCTAAAGC
 TTGTGCAATGATTACCTACCGCTTATTGAAACATTATGATGATCTTCAAAG
 CATGTTGAGTATATATGGAATAAATCTCAAAGATAAAGGCTGATTGTCCAC
 35 TTAACCTTGTTTCATCAGCCATTAAATAAAATCATATATTGTTAAATTTT
 TAAAAATCAGATTGCTCTAAACTCCCGACCGTTACACCTATTAATCGAATCG
 GAATCTCTGGATCTTTTAAAGTCGTTATATAAAGTATAGGCAATATTATAGA
 TATCAGTTTCAGTACGAATAGGATCTCTCAAACCTTTTTTGTCTTGAGATTG
 TTTTCATATTGATAGGTCTTTATTTTACAGTGACTGTTTACCTGATTTT
 40 GAATTTTATTTAAACGTTCTGCTGTTTTTCTGAAAGCTCTCTTATTTTTC
 TTAAAAATAACATCATCATCGTTAACATCTGTAGAAAATGTCCTTTTCAGTCC
 CAACAGATTTACGAACTCTACTCGCCTTTACTTCATTATGGTCTATACCTC
 GAGCTTTATTTACAAAACCTCGTCCTCTTTTACCGAATAAACGTATTAATT
 CAAATTCATCTTTATTATATAAATCTTGACCTGTATATATATGATGTTGAT
 45 GCATTTTTTTCTCGAAGCTTTGCCTACCCCTGGAAAATCTCCAATATCTA
 ATTGCATTAAATATTTTCATGTACATTATTGTAATCAATTACTGTCAAACCAT
 TCGGCTTGTTTCATACCACTCGCTAACCTTTGCTAAAACTTATTATAAGACA
 CGCCAGCTGACGCAGTTAAACGTTTACTTCGTATATATCTCTGCGAATAT
 AATTTGCAATGGTTGATGCTGGTAAATCCGGTCTCACTAAATGTGTAAATAT
 50 CTAATAAAGCTTCATCTAAAGACATGGGTTCTACTAATTCTGTATAACTTC
 TGAATATTTTCATGATTTGACCAGATACCTCTCTATAAGTATCAAAACGGC
 TTGTTACATAATATCCATTGGGGCATAGCTTATGTGCTTGAGTCATAGGCA
 TAGCAGAGTGAACACCATAAGCTCTTGCTTCGTATGATGCCGTAGAACTA
 CGCCTCGATGACTCGCTTTACCGCCAACGATGACAGGTTTTCTTTTAGTT
 55 TAGGATTATCTCTCATTTCAACTTGAGCAAAAAATAATCCATATCTATAT
 GAATAATTCTTCTTTCAGTCATCGACTCACCTCACTATGTCTTCCAAATAA
 TGGCTTTATATTATTATACGCCATCCTTTCTATGTTTACATGTATTGTGCA
 CATGGTAATATTTGTATTGTTTATTTAATTAATGTAATAAAAAAATAGGA
 CTACAATTTTATAGAATAAATTCATAGTCCTATCCTTTTCTACATAAGTAGT

TATGTGAATGTAGTTGCATATTTTAATGAGATTTTTTATTACAATTGAAAC
 GAGTTGAAAAATAAGAAATGAAATTTTAAATCAATTATCAACTTTGATAT
 ATACGCTTGTCACTATAATTAGTGATTATTCATAATAGGTTTGAGTAGGCT
 TTTGATGATAAAGATACAAACCTTCTCCATAGGTTTGATGTATTTGAATT
 5 TGCTTTGTAATGCCAATGCTACTAAAAACACATCTATCATACAATAACTTG
 TATGTATACTAAACATAAAGATTAAAGATGAGAAGGAATATATTTGAATAA
 GTGCTAGTAAGGTCGCACTTATGATAACAAGTGGTGCTATCATGATAATTG
 TAAATTGCCACCGATGAAAACAAGTCTGAGAGAACTGTACAATAATTTTGT
 TTTTATCATATTTTAAAGATGGTTTTTCACCTTTAGAAAAATATAATGAACA
 10 AAATACGATGAATAAACTCGTGCAAAATAACTAGTACTGCGAAACCTACAA
 ATCCAAAAATAAGATTCATGACAAGATTTTGTTCATGATATGCGTTGTTT
 GATAAGCCCATTTATAAGTAAATAGTATTGTAAATAGCGCCAATAC

15 Sequence 3445

step.1003c09.cons.ok
 AATAAATTTTTTGAACGGCTTTGGAATTGATATTCAAATATATCAAAATC
 TAAACCTACCAATAAAATAAAAAATTGCCATACAAATTGCCATAACGACCAG
 TATACTTATTTTTTTATTAGCGCTCATTTGCATAATTTTCTTACCTTTCA
 20 TTAGCAAGATTAGGAAAATAATTGTGCCAAATACACCTATCGTCAATCCAA
 TATTAATTTTCGTAAGGGTAAACAATTATTCTACCAATAATATCAGCAATTA
 AAACAAAAATTGCACCGAGCATAAGCGTGTGTGGTAAAGCATTTTTCAAAT
 GATCGCCTCTATAGATAGATATGATATTCGGTACGATTAAACCTAAAAATG
 GTAATGTTCCAACAGTAACAACAACCTAATGCTGTAGCGTTGCTGTTATGA
 25 ATAATGCTATTTTAATGATTTTTTCATAACTTACACCTAAATTATGACTAA
 AGTCTTTTCCCATACCTGCAATAGTAAATGATTTGCAAATACAAATGCCA
 AAATAAGTAATGGTATTGTGAGATACAAACCTCAAAACGACCACCTCGTTA
 TAACTGCAAAGTTACCAGTTAACCAGTTTCCAATGCCTTTGTAAAGCATTTGG
 TTCTCAACGCTACAAATGTAGTAAACTGGATAAAAATACCACCAATCATAA
 30 TGCTAAAAGTGGAACAAAGATTACATCTTTTACACGGATAAGATTAAATTA
 ATTGGACAAATAAAAAACGTTCCAACAATACTTAGAACACAGCAAATAATA
 ATTTGATTAAAAATGGGACCATTAGGAAAGAACAACAATGACATTAAAAATAC
 CTAATTTTGGCCACTCCATAGTACCAGCAGTCGTTGGGCTTACAAATTTAT
 TTTGCATCATCTGTTCGATAATTAACTCCTGATAAAGCTAGTGAACCTACCCG
 35 AAAGTAGAATACTAACTGTTCTGGGAATTCGACTCGAAAACAAAATATTTA
 TTTGTTTCATCACTTAAATGGAAAAATATCTATTAGAGAGAGCTGACTCACC
 CTATAAAATAAGAGCAATAGTTAAATCACTAATAAAATAAAATAAGGTAT
 AACCTTTAAAAATAAAATTTTCATTAGTTAATCTCCTTGCTAAATATTAAGA
 TGTATAAGGTCAATATACGAAAGTAATGATAATGATAATGATTATCATTTAA
 40 CATCAATATTATAGACTCTATTTCTTTTTTCAATAGTTTTTTCATATTTT
 TTGGGAATTTTATAAATGTTTTATCCAGCAGTAATAATACATCAAAATATT
 TTATAATACCAACTTTTCATAAAAAAGAGATTCTTACTAATTTTAATAGAA
 ATCTCTTTAAAGTACTATATTAATTATCAAAAAACAAAGTCTTCATCACGT
 AGAGGCTCAACATTTAAAGCCAATGTATAGCCATCACCTTTAACAGAGAAG
 45 AAATCGTGGTCTTTGTAGATGTATCTAGTGCATTTTCAATAATAGGGTTA
 AACTCTCTCTCTCGAAATATGGTTCAAAACCTAAATTCGATAATGCTTTA
 TTTCCATTATATCGAACATAATTTAATACATCTTCAGCAAGACCAATATCA
 TCATATAATAAATGTGTATATGAACTTCATTATCATAAAGTTTCATTTAAT
 AATTTGTACATTTCTTGATCAGCTTTTTGTTTCTCACTTTCAGATAACTCA
 50 TTACGTAGACTTTGTGCATCTAAACCTGTGAACACTCCATGTATAGATTCA
 TCTAAAAGTATCTTACGTATAATTTCACTGACGTAGTCATTTTTCTTGT
 CCTGCGAGATATAATGGATAATAGAAGCCAGAATAGAATAGAAATGTTTCT
 AAGAATACACTAGAAACACGAGCAATATATTGATCGTAAATCGATGCTTCT
 TTACCCCAAAGTTTGTGGTAATTTTCTACAATTTTATCTGATTTATATTTT
 55 AAATGTGGCTCTTCAATAACCCAAGTATCCAATAAATAGTTGGTTTCACTA
 GATGGTAATAATGTAGTGAAGATATGAGAATAAATTTTCGCATGGATTTGT
 TCCATCATAGCCATAAATGAATAAACAGCTTTCTTTCTTAAATCAGTAGTA
 TGAAGCATGATTAAATGGCATACCATCATCAGCTTGATGTGTATCTAAACCT
 GTTAAACCTGCAAGCGCTTTTTTAAAGTATTTTTCTCAGAATCTGTTAAT

GTTTTCCAACCTTGCTATATCTTTTGATACTTTAAATTCGTGTTCCACCCAC
 ATTTGTGAGATATTTTGACGCCAGAACATATTAGTCATATCTTCTTGGGTG
 TTCCAATTTACCGCCTTCATTAAAATCCTCCTGTCATATTTATGTATAAAA
 TAAAGAGTTTGAACCAATGTAAATCTATAATTTAAAGAAATACATCAATT
 5 CATCACTCTACTCTAAAATCTTTACAACCTTAATCAATTAAATCGCACAACT
 TGTACATTCTTCCACACTTAATAATTTATTACGTGTATAGTATAAAGATTT
 AAGGCCTTTATGGTGTGCATATACATATAATCTTGAAAGTTACAGTGTGTA
 AATTTCTGAATTAACATAAAGTATCGTTGAAATACCTGGTCTACGTGAGT
 TTGAATTGTAGTACTAAGTCAATAAGTTTCATTTGGTCAGTATTTAAATGC
 10 TGATTTGTAGTACCACATTTGTTCTGGTGATAAAAATGGCATTGGATAAAA
 TGTTTCCGCATTTCCATACGTACGACGTTCAATTTGGTCAACAATAGGCAT
 AACAGAACTTGTGCGATTTTGAACATAAGAAATACTTTGAGTCGGAGCAAT
 AGCTAATCTATAAGCGTGATAAAGTCCATATTGTTCTACTTTATTTTGCAA
 TTCTTTCCAATCATTTGAAGTAGGTATATCGATACCATCAAAAAGTTGGCG
 15 AACCTTTTCAAATTTTGGTTCAAATCTTTGAGATGTATAGAATTTCAAATA
 TTTACCATTTGCATAGTCTGATTGCTCAAAGTCTTGATACTTTTCTCCACG
 CTCTTTTGCATTTTCCATTTGAACGTTTCGATGGAATTAATAGTTTCATAATCAT
 AAAGAATATATTAGCAAAGTCTTTTGCTTCTTCAGATTCATAGCCAATTTT
 ATTTTGTAGTATAGATAACCATGTAAGTTTCATTACTCTAGTCCAACAGAATG
 20 TAGTTCACATATTCGCTTTTTTTTACACCTGGTGCATTTTGAATATTTGCTTC
 ATCACTTACAACGTAAAGAGCATCCATACCTGTGAACACAGAATCTCTGAA
 TTTACCTGATTCCATAACATTCATATATTCAATGAGCCTAAGTTACATGA
 AATATCTCTTTTAATTTTCATCTTCAATTCCATAGTCGTTAATTACTGATGT
 CTCTTGTAAT
 25

Sequence 3446

step.1003c10.cons.ok

TATATTTCTGTTGATTTTCTTTAACTACTGCACCTGCCACAATTCGATGCG
 30 GTTTTTTCTTTAATTTCTTTTAGTAAAGTAATTCCACGTTGTGCGCGTTTAG
 CTTCTTGAAGAACATTAAAATCAATACGCTTCATAGCACCACGTTGTGTAA
 CCATTATTATGGAATCTGAGTCGTTACATCTTCTGTCATAACAACATAGT
 CTTTCATCTTTAAGATTAATTGATTTAACACCAGCTGCTCTTAAGCCTGTAT
 CCGATAATTCATTAGTTGAATAAGTTAATGACATGCCTTTATGGGTTAGAA
 35 CAGTAATTAAGTATGAGTCTAATCGAAGCAGATTAATAAGTTCACTTT
 TGTCTTTAACTTTCATATTTATGAGTGGTTTATTAAACCGAGTAGTTTAA
 ATTTGTGAAGCATCACTTTTCTTAATCATGCCGTTTTTTGTAGCCATAATAT
 AAAAGGCTTCATTTTTTAAATCTTTTTCGTTGTATACATTTACCACTTCTT
 CATCTTCATCTATTGGCACAATTTGTGATATGTGTTGACCAAGCTCTTTCC
 40 AACGGATATCGGCTAATTTATGAACAGGTATAAACAATATCTACCTTTAT
 TTGTAAATACAAGAACAGTATCTTGAGTATTCACGCTTTCATGTTTAAATA
 AACGGTCGCGCTCCTTCAAACCGATTTCAGTCACACCACTTGCGTTAAAAC
 TACGTGTAGATGTACGTTTTATATAGCCATGTTGCGTCAAACCTAAAATCA
 CTTCTTCACTAGGCACCATAACTTCTTTATCAATTTTGATTTCGGAAATTT
 45 CAGCTTCGATTGTAGATAGTCGATCCACTTTAAATTTCTTTTTAATTCAT
 TTAGTTTCATCTTTAATTACTGCTAAAAGTGCCTCATGATTATCTAAGATAT
 TTCTTAATTTCTTTTATTAAAGCTTCTAACTCTTCATGTTCTTTTTTCAAAG
 CTTCAATGTCAGTATTTGTTAATCTATACAGCTGTAACATGACAAATAGCTT
 CTGCTTGAGCTTCAGTAAAGTCATACTCTGCAACTAAATTTATCTTTAGCAT
 50 CTTTTTTATTTTATAGAAATTACGTATCAATGCAATAACTTCATCAAGTATAG
 ATAAAGCTTTTCATTAATCCTTCCACAATATGCATACGTTTTTTCAGCTTGCT
 CTAAGTCATAACGCGTTCTATTTGTAACCACTTCAATTTGATGATTTAAAT
 AACTTTCTATAATTTACGTAATCCCATCAACTTAGGGCGACCTTCACTAA
 TAGCAACCATATTAAAATTTATGAAATTTGTAAATCCGAATCTTATATA
 55 AATAGTTTTTGATTGATTTCGCTATTAGCATCTTTTAAATTCGATTGCAA
 TTCGTAATCCAGTTCTATCAGTCTCATCTCGAATTTCTACAATACCATCAA
 CCTTTTTATCGGCAGTAATTCGTCAATTCTTTTAACTAACTACTTTTAT
 TCACCTCATACGGAATTTTCAGTCACAATTAATTGTTTACGTCCACTTCTTA
 AAGGCTCTTCATCTACTCGTGAACGCACGACAACCTTTCCTTTACCGGTCT

CATACGCTTTTTTTATACCTTCATTCCTTGAATGATACCACCTGTAGGAA
 AGTCAGGCCCTTTGATATATTTTCATCAGTTGATTAATTGTAATATCAGGTT
 GATCGATATACTTCAATGTGCCCTGTATTACTTCGGCGAGGTTATGCGGCG
 GGATATCAGTAGCATATCCTGAAGAAATCCCCGTAGATCCATTAATTAATA
 5 AATTAGGAAATCTCGCTGGTAATACCATTGGTTCCAAAGTTGTGTATCAT
 AGTTTGGAAATAATGATACGTTCCTTATTAATATCCCTTAATAGTTCCTT
 CTGATAATTGACTAAGTTTAGCTTCTGTGTAACGCATAGCAGCTGGAGGAT
 CGTTATCGATACTACCATTATTACCATGCATTTCAATTAGAACATGACGTA
 ACTTCCAATCTTGACTTAAGCGCACCATAGCATCATATACTGAAGAGTCTC
 10 CATGAGGATGATATTGACCTATTACATCACCGACAGTTTTCGCACCTTTTAC
 GGAAATTTTTATCATACGTATTTCCCACTTGAATACATTGCGAATAGTATAC
 GTCGTTGTACTGGTTTTAAACCATCACGTACATCAGGTAATGCACGTTCCTT
 GAATGATATATTTACTATAACGTCCAAATCGGTACCAATAACATCTTCAA
 GTGACAAATCTTGAATTATCTCACTCAATTCGTTTCCTCCTCAATATATTT
 15 TTCAATTCCTAGTATTTGGACTTCTTTATTATCCAAAATGCTTTGATCTTC
 TTGCATAACCAATTCAACGTGTTTTCATCCACTCTCTTCGTGGGGCAAC
 CTTATCCCCCATCAAAGTAGTGACACGTTTGTATGAACGAACCTCATCTTC
 AACTTGAACCTCAATTAATGTCCGAGTTTCTGGATTTCATGGTAGTTTCCCA
 TAATTGTTCTGGATTTCATTTACCAAGACCTTTATAACGCTGTAATATGAA
 20 ACCTTTTCTTAATTGCTTTTGTAAATTTTCTAATTCTTCATCAGTCCAAGC
 GTACTCAACTTTTTTATCTTACCTTTGCCTTTTCTAATTTGTATAAAGG
 CGGTAACGCAATAAAGACACGTCCAGCTTGAACAAGTGGTTTCATATATTT
 AAAGAAAAATGTAAGCAATAAATCTTGAATATGTGCACCATCCGTATCAGC
 ATCTGTATGATAATAATCTGTGTGAATTACTATCCTCAATTTTAAAGTC
 25 AGTACCAACACCAGCACCAATAGTATGAATAATCGTATTAATTTCTTCATT
 TTTAAAAATATCCTCTAAACGTGCCCTTTTCTGTATTAATAACCTTTCCACG
 AAGAGGTAAAATAGCTTGGAAATTTACGGTCGCGTCCCAATTTTGCAGAACCC
 TCCCGCTGAATCACCTTCAACTAGATATAACTCGTTTTCATCAGTATTTTT
 ACTTTGCGCAGGAGTTAACTTACCTGATAACAAATGTATCTTTACGTTTATT
 30 TTTCTTTCCGGAGCGTGCATCTTCTCTAGCTTTACGAGCAGCCTCGCGTGC
 TTGTTGAGCTTAAATTGCTTTTTTAACTAATGATTTAGATAATTGGCCCTT
 TTCTTCTAAGTAATATGGTAATTTTCTGAAACAACAGAGTCTACAGCACT
 CCTTGCTTCTGAAGTGCCAAGTTTGATTTTCGTTTGCCCTTCAAATTGAAG
 AAGTTCTTCTGGTATACGTACTGAAATTATCGCTGTTAAACCTTCGCGTAT
 35 ATCATTACCGTCT

Sequence 3447

step.1003c11.cons.ok

40 CAAATTTACATCGACAAACACTCTTCGATGAAGAAGACGCACATACTTTAA
 TATCCAAAGTTGAAGCAATCAAGCATAAGGTTAATCAAATTAATATAAATG
 TCAATCTAACAACCTTATGATTAGAAAGTTACTTCATCAAATATCGAAAATT
 TGATAAAAAATGTGCAACCAGACATCATCATTGATGGCATGGATAACTTCA
 AAATACGATACCTGATTAATGAGGTTTGTACAAAGTATCAAATCCCATGGG
 45 TTTATGGTGCAGCTGTTGGTAGTAAAGGATCAGTATATGGAATAGATCACC
 AAGGACCATGTCTAAAAATGTTTATTGCAAACAATTCCTGACACAGGGGAAA
 GTTGCGCTATTAATGGCGTAATTCCTTCTGATCAATGATTGCAAGCT
 ATGAAGTAGCAGAGGCCGTACGTTATCTTTCAGGAAAAGGATTTTCAAAGC
 AATTAATCACTAATTGATGCATTTAATATCAATTATAAGTCAATGAATGTAG
 50 ATGCACTCAAAAAATAAAGATTGCCAGTGTGTGAAAAACATGAATATACGT
 TACTAGAAAGCCAACAAGAACGTACTATTGAGGACTTGTGTGGGAATGCTT
 ATTTATTTAGATTCCCACCTAAAGCTTTTAAACAGCTGCCCATTTCCCTG
 GGAATATGGTGAAATCTACTTCCTTTGCCAAATTAATTCAATATCAAACCTT
 ATGAATTCACCTTGTTTAAAGATGGTTCGTATGAATGCATATGGTATACACA
 55 ATGATGAAGAAGCACATCACCTATACAATACGTTGTAAATCCATACGCT
 AATTAAGTAAGATTGTACCCTTTCAACCTGTTTCTTGTAAAATAAATT
 ATACTTAACATCAAGTTTACGACAACCTGGATTAAATTTAAAGGAGAACTA
 ACGTTGTATGGAACAAATTATCACTGATTTTATTAGTAAGTGGGGTTATAC
 AGCGATATTCATTTTAATCTTATTAGAGAACGTATTACCTGTCGTTCCATC

TGAGATTATTTAACTTTTGCAGGCTTATTATCTGTGAAATCACACTTATC
 TATTTGGACATTATTAATCATAGCAACAATTGCTTCATTTCATTGGTTTACT
 CATTTTGTATTATATTTGTAGACTTATCTCAGAAGAGAAATTATATCGTTT
 CGTTGATCGACATGGTAAGTGGATGAAGTTAAAAAGTAAAGATTTGAAACG
 5 GGCAAATGATTGGTTTAAAAAGTATGGTGCCTGGGCTGTATTTTTATGTCG
 TTTTGTCCCAGTACTTCGAGTATTAATTACAATACCTGCTGGCATTAAATCG
 AATGAACGTTATACAGTTTACAACCTTTATCTTTAATAGGTACTACAATTG
 GAATTTTGCCTTAATACTGCTCGGTCGTTTGTCTAGTGACAGTTTGGACGC
 TTTGATGAATGGTATTCATACATATTCACGTATCATGTATGTCATTATTAT
 10 TATTGCACTCATATATTTTGTATACGTTATTTAATGAAACGTCGTCGGAG
 TGTAAATAAATGTTATGCTAAGTTTCCGTTTCTTAACGGTTAACTTAGC
 TTTTTTATTAAATTTAATTTTATGATACGCAATTATGAAAATTAATTGAT
 TTTTCAAGATAATCTATGTAAGCGCATTATTATAAATAATCGTATCGTTAG
 GTCGGAATAATATAATTAGAGGTGACTATGCATGACCCATATTACAGAAAGT
 15 GAAATGAAACAAAAATATCTAGATTACTCTACAAAAAATTTGACAGTGCA
 GAAAAACTTGCTACTGAAATTATTAACCTAGAGTCAATCTTAGAATTACCT
 AAAGGGACTGAACATTTTGTGTAGTGACCTTCATGGTGAATACGAATCTTTC
 CAACATGTTTAAAGAAACGGATCTGGAATGTGCGTGCTAAAATTAATGAT
 ATCTTCAAGATAAATTATCCCAGCAAGAAATCAACGACTTAGCAGCATTA
 20 GTATACTATCCGGAAGAAAAACTAAAAATTAGTTAAAAATAATTCGATTCA
 ATCGGAACATTAAATATTTGGTATATTACAACCATTTCAACGATTAAATGAT
 TTAATTACATATTGCTCATCAAAATATACACGTTCAAAATTTACGCAAGCA
 TTACCTGAACAATACGTTTATATTTATTGAAGAGCTACTTTACAAGAGCAAT
 GAATTTTATATAAATAAGCCTTATTATGAAACATTAGTTAACCATAATTATT
 25 GAATTAGAACAATCAGATGATTAAATCATTGGCCCTTTCCTATACGTACAA
 CGTCTAGTCGTAGACCATCTTCATGTGCTGGGCGATATCTATGACCGTGGT
 CCTAAACCTGATAAGATTATGGATACATTAAATAATTATCATCTCTGTAGAT
 ATCCAATGGGGAAATCATGATGATTATGGATTGGCGCCTATGCTGGTTCA
 AAAGTATGTCTTGCTAACCTTCTACGTATCTGTGCACGTTATGATAATTTA
 30 GATATTATTGAAGATGCATATGGCATCAATCTACGCCCTTTACTTACGCTT
 GCTGAAAAGTATTACGATGCTGAAAACCCAGCGTTTAAACCTAAGAAACGA
 CCAGATAAAGACGTCAGTCTTACAAAACGCGAGAAAGTCAAATCACAAAAA
 TTCATCAAGCAATTGCGATGATTCAATTTAAACTTGAAATGCCTATTATTA
 AGCGTCGTCCTTCTTTTGAGATGGAAGAACGTTTAGTTTGAATAAATAAG
 35 ATTATGATAACAATGAAATCACTATATATAATAAAACATATCCGCTTAAAG
 ATACTTGTTTCCA

Sequence 3448

40 step.1003c12.cons.ok
 TGAGTTCTAATTGGCAATGATTTAAACAGATTGAACAATGTGAGACACAT
 ACTGTTCTTGTATTTCTAAGCTTTCTGCATGATTGTAAAGGCCATTTCTG
 GTTCAATCATCCAGAATTCAATTAAATGACGGCGTGTTTTAGATTTTCTG
 CACGAAAAGTTGGGCCAAATGAAAAAACACGTCCTGAGCCATTGCGGCTG
 45 CTTCCATATACAAC TGCCCACTTTGTGATAAGAATGCATCTTCATCGAAAT
 ATTTTGTATGGAATAAATCACTTGTTCCTCTGGTGCATTGCTGTTAAAA
 TAGGTGGATCAATTTTAGTGAAGCCATTTTCATTTGAAAAACTCATATGTTG
 CACGGATAATTTCAATTTCTTATTTTCATGACAGCATGTTGTTTTTTGAAC
 GTAACCATAAGTGACGATGATCCATTAAAAATCTGTTCCATGATTCTTTG
 50 GTGTAATAGGATAAATCGTGTGCTTCATGTACAAATTCGATTGATTAACTT
 GCATTTCTGAGCCTAAATCAGAACGATTATCTTCTGTAATCGTTCCCTGTGA
 TGTATAAAGATGATTCTTGTAGTTATATCTTTTGTAGTTGAAATGTTCTT
 CATCTACTTCAGATTTTACTACTACTCCTTGCAATAATCCTGTACCATCGC
 GTAATTGTAAAAACGCTATTTTACCACTTGAACGTTTATTAGTTAACCAAG
 55 CACCAATTGTTACTTCTTGGTTAAGATGTTTTTTCGCTTGTTTAATCGTAG
 TTTTCATAACCAAACTCCTATATTTTCTTTTAAATACTTATTATTATTTA
 ACAAAATATGTCTTTAACTTCTACAAGTTGAGTAACTGGAATCATTCGATA
 TTATTTTATTATATTTTCTAAAAACGTTAGCTTTTACTCTCATTTATCGATAG
 ACCTCTGTATTTGATTTCAGTAATTTATTAAACGTTTAATATTACCTTTTC

GTTGTTTAAATGTTCTAGTGATTGAGCAAAAAGCTTTTATATGTGCTAT
 TTAGTAGTCGATCATCAAAATGACACAATAAGTCCTTTGTCGTCTTCGTGTC
 GAATAAGACGTCCAAGACCTTGTCTAAACCTTGTTACAGCATCAGGTAATA
 CATAATCTTTAAAGTTGATGTGAATTCAGAATCCATAAGCCAGTATTTAA
 5 TATTATGTTTATTCATAAATGGTAATTTAGCTATCATTTACACATTTAAGGC
 CATTGGCTTGAAAATCAAAACCTTCAAAAAATGTTGATGTTCCAAGAAAGTA
 TAGATTTATCAAAAGTTATTAATTTGTTGAACAATTTTATAATTTTGGTTTT
 GTTGTGGGTTAAAATGACATAATCTTCTAATCAGGTAATTCATTTAACA
 AATCTTGAACCATATGCATCATTTTATAGCTTGTAATAAGACAAGACATT
 10 TCGACTGTGTCACGTGTGATGTATTCACAATATAGTCTACAATTGAAGCTA
 CATAATCGTCTAAATTTTATAATTGTACGTTTCAACATCATTGGGCACAA
 AAACATTTGTATGGTTGGATGATGTTAAAGGTGTTGAAATTTCAAATGTAT
 TAAAGTCTATATCTTCATTAAACCAATTTTGAAATGCTTTAAATGAGTGAT
 TGAAAGTTAATGTGCCAGATATAAATGTAAGAGACTTAAATTTTTCGAGTA
 15 CTTGTTTCGTCAAGATATCTTTAACATCATAATCCTTGACTAAGAGTCGGA
 TTGTTGATTTTGTGCTAAATTTTAAATTGAAATAAGCTTGATGATGAT
 CTTTTATAGCTTTGTTTCGATAAGTTTAAATTTGTCAATGTAAGTATAATAATT
 GTTTCGCTACAGACTTGATTTGTTTGTGATTCATTCCGTTAAAAATTTCTA
 TCGTTTTATTTAATTTATCTATGATTGCACGTAAATCTTTTAAAATCTCAC
 20 CCGTTTCAAAGTCATAAACGTAATGATACTTATGAATGTCATCATCATAAA
 CGTCTGATGTTTGAATAATATTATAAATTTGTAGTGAATAGTTGCTCATTTA
 AATCATGTAACCTATTGATATTTATTTTCAGTCCAAAAACGTCTATAGGTG
 CAATATCTAGTTTCTCTAAAAATACGTTGTTGCTCAAGTTTGTCTACTGCTT
 TAAGTAGTTTTCATTTTCAATTTTGTGCAATAAGTCCATAATTGATATTTAA
 25 CATCTGAATAATTTAAATCATTAGTAACCTTGATTTAGCGCATAGTCAGGCA
 ATCTATGTGCTTCATCGATGATGCAATCATCAAATAGTTGATATATAGTGT
 TTTCACTGTCTGAATGAATTAAGTGCGCATGATTAGTGATACCAATTTGTA
 TGTTTTGAGCATTTCTTTTATATAATTATAAATAATGGATATCATGACGAA
 CTGGAACGTATGTTTCAATTTTGGTCAACATACATTTTGTCCACCTT
 30 TAAGATTTAATTCCTGTATGTCCCCAGTATTTGTTTCAGTTATCCAAATAA
 GTAACGTGCAATTTAAGAATACTTACTTCATAATTATTTGTATCGTCTTTAA
 GAATTTGGCTGATAAGACCAAGAGATATATAATCATTTTACTTTTGATTA
 ATGACGCGTTAATTTTAAAACTTAAACATCATTTGAGTAATGGTATGTCTT
 TCTCTAATAGCTGACTTTGTAATAATTTGTATTTGTTGAAATCATTACAT
 35 GACGACCGGTCTCAATATTATACATTGTTGCTGCAAGCAGGTAAGCAAGAG
 ACTTTCCACTACCCAAAGGAGCTTCAATCATTTGCTTATCACTATGCATCA
 ACTGATCAAGTATAATTTTCGCTAGGTATAACTGTTGTGGCCGATATGTAA
 GATTTAATGATTGAGTGACATTTTATATAAATCTTTAAGGTACCATCAA
 AATTGACAGCCGGTTTTTTTAAATCAATTTGTTTACGGTAAATAATTGTT
 40 CAAATTGTTCAAATTTGGTTGTTAGGCGGTTTAGTTTGATAATTTCTAACCA
 TTTCAAACAAAATATGATAAAGATCATACTTGAGATTTTACTTAAATAGT
 ACAGTTGTTTGTGTATCTAAATGCAATTGCTCGAACTTCTCAAAGCTT
 TAATCATCAATTTAGCAGTTGTTGTTGCATCTTCATCTGCTCTATGTGCAT
 TATTTAATGGTATATGATGAGATTCAGCCAATGCACCTTAACCTGGTAGCTTT
 45 TGTCTGTAGGGAATGCGATTTTAAACAATTCTAAAGTATCCATTACTCTTT
 TAGGTTTAAATTTGAATATTAC

Sequence 3449

50 step.1003d01.cons.ok
 AGCTTCAGCAATCAAAAATGGTTGGAAAATGAATGACTTTGTAGGACCTGT
 TGATCATGGTGCATATAACCGTATTGAACAGATTAAACGTTGAAGTCACTGA
 GGTTTAAAATACGATATATATAAATACTAGTACATTTTACCCCTCATCTAT
 AGTCTCATTTAATTCCTACAAACATTGTAATTATAATTAAATCAATATCTA
 55 TATAGATGTGGAAAGCCTGGAACAAAAGTTCCAAAATAAGTAAATAAAGA
 CAATTTCTATTGAAATAATATAGGAATTGCTTTTTTATAATTTTTTGAT
 GATTTTCAGCTCGTTGAGCTATTACTTTTCTTATATTAAGCCACGAGGATG
 TCTGAATTCATCGAATTCAGTATCGTGAATTTGTTCTACAATATCATGTAC
 ATATCGTGAAATATCATTCGCGGGGATAAGTACTGAAGTTCTATTTGGTAG

AGTATGTTAAGTCATATTATAATCTTTATACATAAGACACCTCGTTAATTT
 AGTTTAGTGGTATTTATTAAATTTATACGAAAGCACCTTATTTTCAAAGTA
 TTTTAATATAAAATTACATATAAACACAAAGTATTTTGGCGAGACTCTTGA
 5 GGAACAGGACAAGCTGAAGACTACAGGCTGAAGCTGTCACCTAAGAAAAGC
 GAGCCAACAATACGAATTATTTTAAATAAAGATGCCAATAAATGAATTTAT
 GAAAATCCATTTACTGGCTATTTTGCTAGGAATTATGTCTCAGCGTCTCTT
 CTCTTTCCCTTCAATCCATAATAAATATTGTATACATATAAAGTCTTAGAA
 AGTAATCAGTTTATCGTCGTATGTAAATTTGATACTACAGACCCATGATTT
 10 CTTATGATTTTATCTTTTATTATTTAAAGTGTAGAGAGGCAAGGGAGGTATA
 AACACATGATTATCTACAGAAGAAATATAGAAAATGGAACACCCATTTAT
 GAAATCATAACTAAAACCTTTCAAGACAATTACTATAAAGTGTGATGAACT
 TTTAATAAGTATGAAAATCTATCAATTGCTCTCTCTACTAGAGAATGACGTT
 GACAACATGCCGACAAGTTACTCATATCGTTAACGTTATTCTAGACAACATA
 GACATTGCATCAAAACCTAATTGAGATTTAGCACCTTTTGTCTTTCCAGTCCC
 15 AAATAAAATGAGTTTAAAGCATTTAGTCCCTTAATACGTTTAGCTTTAAA
 GACATTCTATCATCCCTAAATTAACACTCACCTTAACAAGTGAGAAACGCT
 AATGTTAGCCAAAACCTTGTCCCTTTTGTAGCTCGGAGAGAACCTTTGAGC
 ATTGCTAACATTGGCGTTTTTAATTATTTAATTTAAAAAAGATTATATT
 TTTATCATTCCAACTTCTAGTAATTGTTAGTTTGGTATCATCATATGTTT
 20 TAAAAACTACAATTTTCATCATTTTCTTTTATTAAAAATGTACTTCTTCAT
 CTATATTTATAGGTTTTAAGTACTTAATTTTGTAAATTCATCCAATTCACAT
 TAAGTTCAGAAAATACTTTTTACATATGAGTTGTCTGGTACAATGGAAC
 GATGGATCGGATTCTCATCAATTAACATATAGTTAAATATTTCATTCACCATTT
 CTATTGTAAATTTTCATAATTTATCTTACCTCGATAAAATATTTGTTTTATAT
 25 ACATACAATGTTTTTATCTTTATTGATTTTCGAGTCCATATGTATATTTAA
 TGATATTTTAAATTAATTTTTGTGAAACCTTATGCAATGTCGCTTCATACG
 TACAATCGACTTGAAGTTGATGTTGTGTCTTAATGTGTGTCTCTTTAATA
 TGATAGGCTTCTTTGAAAAAGATTGAAATAACTCAAATTTGTGGCCATAACT
 TGGCACACATAAGAGTAGGAACATAAATGGTCATACTTTATATTAAAGCAATT
 30 CGCAATAATTTTGAACCTTCTTCATTTGTAAATTGAACCTTCTCTAATTTCCA
 TACCATCTTTCAAATAAAGCTGCATTTCTATCCCTCCACCTATTCCCAT
 GTTGCTATAGTTCTAAATTTGATGTTTCATATAAAATAAACGTGTGACTAAC
 GCTGCTCCACTTGCACCATATGGATGCCCTGTAGCAATTGCTCCTCCCAA
 CAATTCAACTTATTCAAAGGAATGTTAAGCTGTTGTTGGCTCGCAATAACT
 35 TGAGAGCTAAATGCTTCATTTAATTTCTACTGCATTTATATCATTTATAGTT
 AATCGTTCTTGAGCTAATAAATGATTTACTGCTGGCACTGGACCTACTCCT
 AAATACGTGGTTGAACACCTACAGTTGCACATTTACAAAACCTTAATCCCT
 TCTGTGAATCCTAATTGACGTGCCCGATTTTCTTCCATAACAATCAGTAAT
 ACTGCACCATCATTTTTCTTACAACATTTTCTACTGTGACTGTTCCCTTCA
 40 TTTAAAAGTGGTTTAAAGTCTGCCAAGTGTTCTGAGAGTAAGTTGAGGTTTA
 ATACTTTCATCTTGATTAAAAATTTACCTTTCACTTTGAACGGTAAAAAT
 TCTTGGGAAATATTACCGTTATTCATATTTTTTGATGCCAATGATGACTG
 CGATACGCAAAAGTCATCTTGTTCATTTCTACTGATATGATATTTCTTCGCT
 ACATTTTCGGCTGCTTCAATCATTTGAAGGGTCTTCTCCTTCTCTTGCAAAA
 45 GGCGCCCGTTCAAAAAATTTGTGAAACTCAGATTCATAAACTGACTGCGGA
 CGTTTGATTTTCCAAGGTGCTCTACTGGTACTCTCAACACCACCTGCAATA
 TATATTGTTCCAGCACCACCTTTGTACCATCCTACAGGCTTGTATAACGGCT
 TCAAGACCTGAGCCACATTTGACGATCAATTGTTATACCAGGTATTTAAAA
 TCTAATCCCGCTTCAAGTAATGATTTTCTAGCTAAATTCCCCCATTTACCT
 50 ACCGTATTACCTA

Sequence 3450

step.1003d03.cons.ok

55 AATGCTACGGCTCAACTTCGTCAAAACCATTCAAAGAGCCAAAGTGATTGG
 AATCATCAACAAAACCAACAACAGAAAGATGCTTGGGATCCTGAACAAGAA
 ATCAACAAACAAAAAAGGACGATCCATATATATATTAATAGGTTTAGATT
 GATGTTTGACTATTAAATGTTGAATGCAGTCCATAAAAAATACGCCCTATCA
 GTAAAAATAATTGATAAGGCGTATTTTTTATAAAGTAAGAAAGGGACTATA

CAACTTGAAGTTAACTATATGAATAAATAGTGATAGAAAATGACTTTTAA
TTCGTCATTTCAAAATAGAGGATACATTTTCTTTGTTATTCATTTAATCAT
AATTCTTGGAAATGTTTGAATGATAAGATATACGTTAAGTCCACTTAAGACA
ATTATGAGTAACATATCAAAAATATGTACACCAAAGTCCTTAGCATTTTGT
5 TGTAAGTTGTAATTTGTTTATTAGCAATTTTCATCTTTTATATTAATAAATA
TCAATTGTGGGAACGTTATGATCTAAAGTTGCAAAGGTTAGATCAGGCTT
CTGAGTTTACGATTTTGTATTC TAAGTCCTTCAAACGCTTGAGGAGAAGTG
ACTTCATGAATGAGATGTAAATCAATGTATAATAATTGTGGTTCACCTTCT
TTTCCATGAAGCAGATGCTTTTCCATACCTTATCAAACAGTGTTTGACCC
10 ATAATTTCTCCTCCTCTTATAGATATTTTCTTTAAGCAATTTAAAAAT
TCTGAAGTTTCGATATTGTCCACCTAAATCTGCAGTTGTCTTATTAGATTGA
ATAAACGAGTAAACAATTGACTCAAGTTTCGTTAGCAGCATCATTTTGATTT
AAACTTTCTCTTAAGCAAAGTGCTAAAGATAGAACCATAACCAATGGATTC
GCTTTATCTTCATTAGCTATATCTGGTGCTGAACCATGAATTGGTTCATAA
15 AGACGTGTACCTGTTTGACCAAACTAGCTGATGGAGAAAGACCTAGAGAC
CCTGGTATAACAGATGCTTTCATCACTTAAATATCTCCGAAAAGATTCTCT
GTTACAATCACATCAAATTCGCTAGGTTGAGTGATCAGATGCATGCTACAA
GCATCAACTAGCATATGATTAACTTCTACTTCTGGATAATCCTTTTACAA
TCGTTTACTATTTGTGCGCCATAACTTACTAGAGGATAGAACGTTTTCCTTA
20 TCAACAGAAAGTTAATTTTACGCTACGATTTGCAAGATTAAAAGCGATT
CTCACAATACGTTCAATTTCTTGACTTGAATATTTAATGAGCTAAAGCT
TCGGTTTTTTAACATAGCTTGGTTCACCAAAATAAATCCCGCTCGTTAAT
TCTCTTACTATAACTAAATCTGTGCTTCTACAATATCTTGTTTTAGTGGT
GAAAGATGACTAGCCCCCTTAGTTACAAATGTAGGTGCAATATTAGCAAAT
25 AAATTTAAGGATTTCTCAGTTTAAACAAACCGTGTCTGGACGATTATTT
GGGTCTGTCCATTTAGGCCCTCCAATCGCGCCTAATAAAATAGCATCTGCA
TTTTTACATGATTGTAATGTTTCGTTAGTTAAAGGTGTACCATAATAATCT
ATTGAAACACCTCCAAAATGATGACTTTCCAAATGGTATTCAAATGATAT
TTTTCACTTATTAATTTTAAAGTTCTAATGTACCAGATAGAATCTCTGGT
30 CCTATACCATCTCCAGGTAAGGCAACTATTTTATAACTCATGTACGTATAC
CTTCTTTCAACTCATATTTCTGAAATATATTTAGCATGAGCATCGATATATG
CTTTACATGAAGCTTTTAAATATATCGTGGTCTATACCTATGCCTGTCACTT
CAATATGATTAAATAATGATTCTGATCATGTACTTCTGCTTGAGCATCAGTAC
CTTCTGTTACAGAATCAATACGATAATCAATTAATTTCTGCGTCTTTCTTGA
35 AAATTCGGTCAACAGCATTATAAATTGCAACAATTGAACCCGTTCCAATAC
TTGAATCTTGTTTAACTTGTTCGGTTTCTTTCCCTTTATAACTACTACTGCAC
TTTGTAGACCTTTAGATACGTATTGAAGTTGTAAGTTATCAAGTTGAAAAA
TAGCATTTGTGTTTCATGTTCCGAGCCATGTATAATCGCATGAATATCTCTAT
CGGATACATTTTTTTCTTATCGGCAATTTCTTTAAATGTTTAAATAATG
40 TAACTTGATCTTCCAATTTAATTTTCATATCCCAGAGCTTTAAGCTTTTCGG
CAAATGCGATGTTTACCAGACAATTTACCTAGTGGCAATTCTGTTGTATTCA
CACCTACAAGTTGAGGTGTCATGATTTCATAGGTTTCACGATGTTTAAGGA
CACCGTCTTGGTGAATTCGGATTTCATGACTAAATGCATTTTGACCGACTA
TAGCTTTATTTCTAGGTACACGGATACCAGCATATCTTGAAATTAAGTCAG
45 ATGTTTTCTTAGTTTCTTCAAGGTAAATTTGAGATTCAAGACCATAGTGGT
CCTTCTTACATATAAAGCCAAAGCGACTTCTTCAAGTGAGGCATTTCTCTG
CTCTTTACCAATACCGTTACCGGTACCTTCAATACGTCTAGCTCCACCTT
CAATAGCTGCTAAACTATTAGCAACTGCCATTCGAGATCATCATGACAAT
GTGCACTAAAGATAATTTTGAATTTGACTTAACGGCTGTGTTAATTGTT
50 TAAAAATTTGCCCATACTCTGTAGGATAACTAAATCCAACGTGATCAGGGA
TGTTGATAAATGTGGCTCCGCTTAATCGCAGTTTGAACACATTCAATTA
AAAATGGAATTTTCAGTTCTTGTGTCATCCTCTGGAGAGAATTGTACGACTT
CAAAAAATGTTTTGCATAAGAAACGTGTCTTTTATTGATGTTAAACTT
CATCTTGAGTCATTTTTAATTTATGTTCTAAATGAATAGGGGAGGTTGCAA
55 TGAATACATGTACTTGAGGTTTTACAGCTTCTTAGTGGCTTCATATACAG
CATCAATATCAGATTTTTTACATCTAGCTAAAC

Sequence 3451

step.1003d04.cons.ok
GAGAAGCAACAGTAATATCCACAACCTTTAGTGTGTATCTATTACGTTTCG
CTATTGTCAATTGCCGAAACAATTAGAAATGCAAGATCAATTTTTCTATTTTT
ATTTAACAGTTGTCATTTTCATGCTTAATTGCAGCAATGATTATGCCAAGAA
5 TTTGGCCACTTAAAAATATTCCTGACGAATATGCTAAAGAAGTAAGTGAAG
AGGCTCGTAATGAACAGCTACCAGAAGGCAAAACAGCATTAATAATATGGTT
TTGATTTAGCAACTGAAGTTGGAATTAATTCGCCAGGGTTAAAGAATTTT
TAATTTACAGGTTTTAAACAGTTGTAGATATGTGGTTTGTAAATTTTACCAG
TTGTTATGAGTATAGGAACAATAGCTACCATTATTGCTAACTACACGCCCTG
10 TTTTTGAAATTAATAGGAAAACCAATTTGTTCAGTACTAGAATTGTTACAAA
TTCCAGAAGCAGCATGAAGCATCACAAACAATTTTAATTGGGTTTGCCGATA
TGTTCTTACCCTCAATTCTTATTGAAGGGGTTCAAAATGATGTAACACGTT
TTGTAATTGGAGCATTGAGTATCTCACAACCTGTGTATTTATCTGAAGTAG
GCGGCGTGATTCTTGGTTCTAAATTCAGTTAGTATAAGTAAATTTATTTA
15 TGATTTTTTTTAATTCGTACTATCATTACGCTTCCAATAATTGCTTTATTAG
CGCATTTATTTATCGGATAAAAGATATACTCATTTTTTAATTTAGTGAGAGT
GGGGGGAAGAAAATCATAACAACAGTAATATAAAATTAATAATGGTTGAAAT
AAAAATTCATATATCAAACCTGTCTCAATTTTTTATAATTATAAAAATTA
GACAATAAATCATATACGTATTCCTAGCCCACTCCCTTAGATAATTTATCT
20 TCTGTTCAAAGAGATAAAACATTTGATTTAACAATTTTCGCTTTTAATGAC
TGGTTATTTAACATTTATAGTCCCTCCAATTATTTACTTAGACATTTATATT
ATATAGAAAATAGCGCTTTTTTACTATGATAGATGTACAATCAAAATTAAT
ACAATTTTATAAATTCATTTAAACCTGCATATACGATAGCAATTCAGGGCA
AGAACTATTTTATTAGGTAGCCCTCTTTATTTACTTGTCTTGACGCGCTA
25 ATGCTAAACTATCTTCAGGGATATTATCTGTAATTGTGGATCCAGCTGCAA
TAAGAGTATGATTACCTACTGTTACAGGTGCTATAAGGTTTCGTATTACATC
CTATAAAAGCATCTTTACCTACTATTTGTTTTAAATTTGTTAGCGCCGTCAT
AGTTTACAGTGATAGAACCACATCCGATATTTGTACGCTCACCTATTTTCAG
CATCTCCAATGTAGCTCAAGTGTGATACTTTAGCACCATCTTTAATATCTG
30 CTTTCTTCACTTCAACAAAGTTTCCCTACTTTAACTTCAGAACCTAAATTAG
ATCCTGGGCGTAGTTGAGCGAAAGGTCCAACCTGTTGTGTTCTCTCCAACAA
TAGAGTCATTGATAACCGATTGTTTTATGTTAGCATTCGAATGAATTGTGC
TATTATTAATTTTCAGAGTATTGGCCTATCCACACATCTTCTTCAATCGTTG
TATGACCTCCGATGCGCACGCCCGGTTCAATAGTTGTATCTATTCCAATTT
35 TCACATCTGTTCCAATAAATGTTGAAGTGGATCAATGATTGTCACACCAT
TTTCCATATGATAGCGATTGATACGTTGTTGCAAAGCCTTTTCAGCTTCAC
TTAACATCAAACGATCATTAAACACCAATGATTTTCATCAAAATCTTCGGTAC
AATAAACTTCAGCTTTACCCTCCGCTTTTTAAATTTAAAGACAAAACATCAG
GTAAATAATATCTCTCTTGAGCATTTATCATTTTTAACTTGTCTAACTTCT
40 CAAATAGTACTCGATTATTAAAGGCAAAAATACCTGAACATAATTTCTTTAA
TCGCACGTTCTGAGTCATTTGCGTCTTTCTCTTCAACAATACGCTCTAATA
TTCCATTATGATTTCTAATAATTCGTCCATAACCATAAGGATTGATAGTAG
AAGCAGATAATACAGTAACATGTGATTGTGTACTTTCATGATGTTCAATAA
GTGATTGTAAGTTTGGTATGTAATAAGTGGTGTATCTCCACATACTACTA
45 GAGTAGTTCCCTCTTTATCTGCTAAATGTTTCATGTGCCATTTTCACAGCAT
GAGCTGTTCCAAGTTGTTTATCCTGAAAACATATAATGATTGATTACCCA
ATGTATCTTTTCACACTCTCAGCGCCATGACCAATAATAGTTACAATTTGAT
CAACGCCAGCTTGTTTTACGTTGTTAAGCACATGTTCAACCATTTGGTTTGC
CAGCAACCTCATAGGAGCACTTTATATTTCTTTGATTTCATTCTTTGTGCCCT
50 TACCTGCTGCCAGATAAATCGCATGTCTTTGCATGAATGCTAACCCCTCCAT
TAAATTTTACACTTCATAGTTATTATAGTTGAATTAACAAAGAGAGTTTCAA
CTTATAAAAGGTTTTGTAAATCATTCATTAAAGGTTTGATAATTTTAGAAA
TTCAAATAAGTTGTTATGCTTTAATAAGGAAAACGAGGTTCATCAACCCAA
CAGTTACATTACATATTCAACCTCTTTATTAATCATTTGGTATTTTATTTCA
55 GTGTCTAAATCATCATATTATCAATCATAAATTTATGACGATAAAGTTATAA
TCATACAGACATATTAAATACGTCTATCATAAGTAAATTTTATAGATGAA
TTGTTGTGAGAGAAACATATCAGTATTTTACTTCTATAATTCTTATTTATC
GTTAAATGTTTATTTTCGTTTTTCGTTTATGAAATACAAAGTACATTAATAA
TAACGGAGGGACTACAACCTCAATAAGAAGTCGCAGTCCCTCCCGAACAAT

Sequence 3452

5 step.1003d06.cons.ok

GTAGCCTAATGGCACATTATAATTATATTCAGAAATTTTGATTTTCCATT
CTTATATACTTTTTCAACCACTGCTACGTGCCCATAATAGCCTCTCGTTGA
TTGCATAACACGCGTATTTTCGAGGCTTATGTCCAGTTGCATAACCAGATCG
ACGTGCATTGCTATACCAATTCTTAGCATTACCCCATCGATTTGATACAGG
10 TTTACCTAATTGAGCACGACGTTTAAATGCCACCAAGTGCATTGTTTTT
ACCATAATAATTATGTGCAGCATGGGAAATATGCGTTTGTTCAATAGCTTG
ATATGAGAGTGTTCAGTAAATGTAATCATTGTAGCGATAATAATACGAGA
TAATAATTTTGAATTTTCATTTCTTTCTCCCAAAAAATGTATTTATTAA
TGACAGTAATAATAATAGTATTTAAAAATGCGCTATACAATATAAGAAATG
15 AATTATCAAAGTAAATGTAACATTAATGAAATATACATATTTTATCGTAA
TAATCGTTTGAGCCGACTTTTATTTTAACATATTTTCGTTTTTTAAATAC
TCACTCAGGTACGAGATGATTAACTCCATGTATATAAGTATGCGTAAATGT
TAAAGTCATGTCATAACATCAATTTTACTTTTAGTGAATTGGTTGAAAAA
GCTAGTGCCCGTTTGATACCTCGTCACTGGTATGTTTCCGGACATGGATGA
20 TTTTCGAAAATATGTCAAGATTCTTTTGGCTCGTCACGTCTTTAGGGAAA
TTGATATCTTGATTAATCCAATCAACCAATTCACCTAATGGCGTGTCAATCA
CCTAGGAAATTTGCATAAAATCGTAAAAGCTCAATATCTCACCTCATAAT
ATATTTTATGAGCCGACTTTTATTTTAACATATTTTCGTTTTTTAAATAC
TTGAAAACTAAATAAAATGTTTTGAATTATAGTTTTTAAATATATTAATC
25 GCTAGAAAGGTCAATTTTGTGTATAATTAACTTACTATTGAATTATATC
AAAAC TAGGGGATATGCATGGATATAAAACACATGAAATACTTTGTGAA
GTTGTCAAACAAGGTGGTATGACAAACGCTTCAAATCCTTGTATATTCG
CAACCAACTATCAGTAAAGCAATTAAGGATATTGAAGCAGAGATGGCTGTC
CCTTTATTTGACCGGAGTAAAAGAAGTTTAGTACTTACTGATGCAGGTAAA
30 ATTTTTTTCAAGAAATGTCAAGAAATCATCGCACTATATGATAATTGCCC
ACTGAAATTAATAGTTTGTATGGTTTAGAAACAGGTCATATCACTATTAGT
ATGTCTGCAGTGATGAGCATGCGTAAATTTATTGGCGTATTAGGAGACTTT
CATCAACTTTATCCGAATATTACGTACAACTTAATCGAAAGTGGTGGTAAG
ACGACTGAAAACCTTATACCTAATGATGAAGTGGATATTGGTGTGACAACA
35 TTGCCAGTAGATCATCAAAAATTTGAATGTATATCTTTAAACAAGAAGAA
CTGACTGTAGTTTTAAATAAAGAACATCCTTTAGCACAAAAATCTTCTATT
AAAATGGAAGAATTAGCTGATGAGAACTTCATTTTATTTAATGAAGATTTC
TATCTCAACGATAAAAATTATTGAAAATGCGAAGAATGCTGGATTTCGTGCCG
AACATGGCCTCACAATCTCAATGGAATGTGATTGAAAATCTTGTCATT
40 AATCAATTAGGTATTTCCATATTGCCAGCCACTATAGCACAACTTACTTAAT
GATGACGTCAAAATTTGACATTTGGAAAATGCACATACAACCTTGGGAGCTT
GGTGTCTGTTTGGAAAAAGATAAACGTTTAAGTCATGCTACAAATAAATGG
ATAGAATTTTGAAGAAAGATTATCCGAAGAAATAATATAAATATAATAAC
AAACAATGCATTTAATAAAAGAGCGAAGTTGTAAAGATGTGAAATAATAT
45 GTTTTTACAAAGTCATATTACAACGAAAACGTAGGGTTTTGAATTATAAAT
TAATTCGCCCCACGTTTTTTTACTATCAATTAGTTCTCAGCACTATTAAT
TTGCTAGCGTTTTTCATAGATTTAAGGAATATAAAATATCATGAATAAATAT
TTTTAATATATAATAGCAAACTATAATAGAAAATGTCACATAAAAAATA
CAGTAACAATTTGGACACATTTTGTAGAAAGGGTGTGTTGATGGAAAAAG
50 CAAAATTTGTCAATTAAAGTTAATACTTCAACTTGCCCTTATCATGCTTATTA
CTTTTATAGGCACAGAAGTTCAAAAATTACTTCATATACCTCTAGCAGGTA
GTATCGTAGGGCTTATGCTTTTTTCTTATTTGTTACAATTTAAAAATTGTAC
CTGAATCATGGATTAATGTAGGAGCAGACTTTTTTACTTAAACAAATGGTTT
TCTTCTTTATCCCATCAGTGGTAGGAATTATGGATGTTGCATCTAATATCA
55 CGATGAATTATATATTCTTTATTGTTATTATAATTGGTACATGCCTTG
TAGCACTATCATCAGGTTATATCGCTGAAAAATGCTAGAAAAAGCAATA
CAGTAAAGGAACCTGATCACTCATGAATGAATACTTACAAGCAGTTCTAAT
GATTTTGTAACTATTGTTCTGTACTATGTTTCCAAAAAGATTCAAGATAA
ATACAATAATCCACTATTTAAATCCAGCTCTTATTGCATCAATTGCAATAAT

5 TATTGTTTATTGGTTTGCAGGTAAGCTATAAGGGGTATATGAAAGGTGG
TACCTGGATTAAACCATGTTTAAACGCTACAGTTGTATGTCCTGCATACCC
ACTTTATCAAAATAAAAAGAAAATAAAAAATATTTAAACAATTATTTTCAC
AAGCGTGTGACTGGTGTAGTTCTCAATTTTGTGTAGTATTTACAACGTT
GAAAATCTTTGGTTATTCTAAAGACACAATTGTTACCCTGTTACCTA

Sequence 3453

step.1003d08.cons.ok

10 GTTATCCTCTTGTAAAAACATTTTCAATTTCTAATATTTTAAAAATTC
ACTAGCTGCATATTTTAAATGAAGATTCATCAATGTCAAAATGAGGATTATG
ATGAGGTGCAGTAATGCCCTTGTCTTTATTACCACAACCAGTTAAGAAGAA
TGCACCAGGTCGTACTTTAAGGTAATGTGAAAAGTCCTCACCATCATCAT
15 TAAGTCGACTCATTAAATCTTAAATGTAAATCATTAGCTGCTTGTGAC
TACATCATATGATTGTTGATGATTGTGTACTGGTAAATAACCTTTAATATA
TTCCAATTCATATGTAATATCGTTTGATAAAGATAAACCTTGTAAGAGCTT
ATCCATTTTATTTTGAACATGACTTTGTAATTTTGTGTCAAAAGTACGTAC
AGTACCTTTACAAAAGCTGTATCTGGAATAACACTATCTGTTGATCCTGC
TTGAATCATTCGGAAGTAAGAACAGCTTCCTTTACTGGATCAATTGTTG
20 AGAAATTTATTTTGGGCACTTAAAATAAACTCAGCCATAATGACAATAGG
GTCTATTGTTTCGTGTGGTTTTCACCGTGACCCTTTTCCATATATAGT
CACACTAAATTCATCTGGTGAAGCCATTATTGCTCCAGGTCTAGAATAGAT
TGTCCCAGATGGATAAACCCTCCATAAGTGTGTGCCATATATTTTATCGAC
ATTCGTGTAGACAGCCATCATCAATCATCTCTTGAGAACCACCTGGCATAAT
25 TTCCCTCACCATATTGGAATATAAAAAACAACATTACCTTTCAATAAATGACG
ATGCTCATGAACAATTTTCAGTACTCCAAGCAAAATAGCTGTATGTCCGTC
ATGTCCACAAGCATGCATGCACCCCTTTATTTTGAACGATAGGGTACATC
ATTTAAATTCCTGAACAGGTAGTGCATCGAAATCTGCTCGTAATGCAATCGT
TGGTCCATTTGAATCAGATCCCTTTAAAAGTTGCTTTTATACCATTACGTCC
30 AACTGGTGTCTTATGGTGCAATGCTAATTGGCTCAATTGATTTACAATAAA
ATCATGCGTACGTTTTTCTCAAAAGAAAGTTCTGGGTATTGATGCAATAA
TCGTGTAATTCACCATCTCTTTTCTTTATTACTTGCTAATTGAAACCA
ATCAACACATGTATCACTTCTTCCATCTTAATTATTGATATTAGTATATC
ATTATTGTTTATAGCTTAAATAGACTTTGTTTTGTATTAAAGATATCACACG
35 TTTTTTACACTATCTATTAAAGATTGGAACATTAAAATGATTAAATATAA
ATGAATAGCCCCATAAAAAACAACACAAATCATAATACCTCACTTAAAAT
TAATAAGAAAAATCGTCCATAAAATTTTGTATTAACTGTTATTATAGA
AGATAAAGTTTATGTAAAAGTATTCAAAAATATGTAGAATACTAAAATACT
AGGTAGTGATGTGAATTAGGCGTATTCTAAGAATTTAATTGGTATACGAC
40 ATTTTAAATATTATATGAAGTTCTGGCTAAAGGAATATTAGAATTAGTAATT
ATAGTTTATTAGAAGTGGAAAAAATTAAAAAGTAAGAAAAATTATTATAAAA
TAATACCTTAACTTATATGTAGAAATATAAAATATTATGATATAGTTTGT
TATAATTTTGTATGTTTACAGTTAATCACATTAAAATAGAGTTATCAATAA
TATAAGTTATGGGGTATTTAAGTGGAGGAAATAAGACCAATAACCTATCAA
45 GATAAAGAGGCATACCTATTATTATATTCAAGAGTGGTATGAAAATGAGGAA
AAAGTGGTACCAGGGAATACAGATATTGCCAATTATAGTTCATTAAACAAT
ATGGTTGATCGGCTTAATTGTAGTGAAGTTGATGAGGGTTTGTACCGACT
ACAACACTATTTTATTTTAAAGATTCAATTATTATAGGTGCTGTTGATATC
AGACATCAATTAAATGATAAACTATCTAATATTGGTGGTCAATGTTGGGATAC
50 GGTGTAGCTAAATCTTATAGAGGGAAGGTTACGCTACTATCCTTCTAGAA
AAGGCTTTAGTGAACCTTAAGACATTAAATGTAGAGGTGCTACTTATGACT
TGTAATCCACTTAATTTTGGCTTCTCAAACAGTGATGAAGAAATGTGGTGGA
TATCAAATGAATCCTATATTAAAAAAATGGTAAACCTGTTTCATCGATAT
CATATACCTAATACAAAAATCATGATTATTAAGAATATATTGTGATAAA
55 GCAAAATATATGATTTTATTTTAAATGATTAGAAGAAAGAGCTTTTAA
AAGTCAATATTATTTGAGATAGGAGCTTAGACTTTTAAAGTTTATTTAAGCA
AAAGAATTTTAAATGGAGACTTACGATTTAACACATCTTGTATATTAAACT
TTGTGAAAAACTTAACCATCTAAATAATCCTATATAAATACATTAAACGCGT
GGGGGGGATTCAATAAATGAAATATATAGATGAGCAAAACACAAGCTCAAT

TATTAGACATGAATGAAGTCATTTTAGAAGTAGAAAAAGCTTTGCAAGCTT
TTTCAGAGAATAAGACGATAACGCCATTAAGATATGTTTGCCATTTAATG
AGCAAAATCGTTATTTAGTGATGCCAGCATTATCAGATGAATTAAATATCG
TTGGACTTAAACAGTCTCATTTCACCTGAAAATTCAAAAAAGGGAAG
5 CGACTATTACTGGATCAGTTATTTAAGTGAATGAAACAGGAGAAACAT
TGTCTATATTAGATGGTGGTTTTCTTACTAAAGTAAGAACTGGTGCAATTT
CAGGTGTAGCTACTAAATATCTAGCAAAAGAAAACGCTAAAACACTTAGTG
TAATAGGGGAGGTGTACAAGCTGAAGGTTTAATTGAAGCGATACCTTGCTG
TTAGAGATATTGAAAAAATTCACATCGCAAGTAGAACGTTTCGAAAAAGCAG
10 AAAAATTTGCTCAAAATATACGAAATCGATTTAATATTAAAGTGAGTGAT
TTAGATTTCGGCAGATGAAGCGATAGACAGTGCAGATATTGTAGTTACAGCAA
CAAATGCAAATCAGCCCGTTTATACTCATTCTTTACATCCAGGCGTGCAAT
TAAATGCAGTCGGATCCTTTAAACCAGATATGCAAGAAATACCTTCAGAAA
CAATGCTTGTGTCTAATAAAATTGTTGTTGAATCTATGGAAGCAGCTTTAG
15 AAGAAACAGGTGATTTAAAAATTCCTCAAGCAGAAGGAATATTAACATAAAA
ATATGCTACATAGTGAATTAGGCGACATTATTCT

Sequence 3454

20 step.1003d09.cons.ok
GTGAGGGTGCTTCTGGTAAAGATGCTATCACTAAAGCGCATGAGCTAAAAC
CAGATTTAATATTAATGGATTTACTTATGGACGATATGGATGGTGTAGAAG
CAACTACTGAAATAAAAAAGATTTACCTCAAATTTAAAGTAGTCATGTTAA
CAAGCTTTATAGAGGATAAAGAAGTTTATCGTGCACCTTGATTCTGGAGTAG
25 ATAGTTATATTTTAAAGACAACAAGTGCAAGTGATATAGCTGACGCTGTGC
GTAAAACGTATGAAGGTGAATCAGTATTTGAACCAGAAAGTGTAGTAAAAA
TGCGCAATCGTATGAAAAAACGTGCCGAGCTTTATGAAATGTTGACAGAAA
GAGAAATGGAGATCCTATTACTTATAGCTAAAGGATACTCTAACCAAGAGA
TTGCAAGCGCCTCTCATATCACCATCAAAACAGTAAAACTCATGTAAAGTA
30 ACATACTAAGTAAATTAGAAGTACAAGATCGAACACAAGCAGTAATATATG
CGTTCCAGCATAAATTTAATTCAATAAATTATTAAGGCGAAAGTAAAGATA
CATCTATCGATACTTTTCGCCCTTCAATATTATCAGTCGCTTGTGTGATGTT
CTTCTTTAGATTGATGATTGATATTATACGTGTCATCTTCATTAGAGATAT
CTACATCTTTGCTATCTTCATAGTATTCATAAGTCGTATTTTCGTTATAGT
35 GATTTTCAATGTTATCATCATGTTTAATAGCGGCTTCTTCAGGCGTGTGAC
CAGCTATGACTTTTCTTTGGTGAATAATTGCATTGATTTTCAGCACCAGATAA
TAATAATAAGCTTGTGATATACAACCATAAGAATAAAATGATAATACCAG
CTAAACTTCCGTATGTTTTTCGAATAGTTACTAAAGTTTGAAATATAGTAAC
CGAATGCAAAGGACCCTAGTAACCAAATAATGGAAGTGAAATAGCGCCAG
40 GAATAACAGAACGTAATTTTGTTTTAACATTAGGTGCAACTGAATAAAGTA
CAGTAAATATGATGAAAATAATAATCAATGGAATAACAATTTCGTACTAAAT
TAAAAATCCATTCAATTTGATTATCAATACCTAGTGGTCCAAATAAAAAAT
TATTAATGACTGGACCTAGTGTATAAGTACTACAGCAACAACAAATACTG
CACCTAAAACAAGTGATATAGAATACTTAATAATTAACTACTACACCGT
45 TTCGAGAATCCTCAACGTCATAAGCAACATTAAATGAGTTAATAATAGCGG
ACATTCCATTTGATGCTGACCAAATAGCTAAAATTAACCAACTGACAAAA
TACCTCCACTGGCAGTGTCAGAGATATCTTTAACAATTCACCAACGATGC
TAGCTGTTTCTTGATCAGGGACATATTGACTTACTTTTGTATTAACTGAT
TAGCATCAATCGTGATAAATTGACCAAGTAGCGTAAGTAAAAAATTAGCA
50 TTGGGAAAAGTGCTAATACAAAATGATATGTCATCTGTGCTGCTAAACCTG
CAGCGTCATCTTTACCTATCCTATAAATCAGGTAAAGAAAAGAAATTAGAGT
TTTCGGTGATTTTGTGCTGGTTTATTCAGCCGGGAAACAAAAAATACTTGAT
TCTTTTTTTTAGGTTTCTTAGATTGGAATTCTTGAGGTTCTATATATGTAC
GGTCAACTTCTATTTTTTTTATCTTTTTATGCTCTTTATCTTCTATTGAAT
55 TAAGATATTTAGAAGTTGTTTCTCTTTTTTTTGACATAACAATCTCCTTTG
CTTTTTATGTTTTACTTTTTAAAAATTGATATCCCGTATTTGCGGTATAGTA
AAACAAATATGGTTTTATTTATATATGATAAATCCAATACCCTTTAAAAAG
TAATGTATTTTAACTCAATGATTATTTTGGCCTAAATTAATAGAAAG
GCAATAGTTTTTAAATATTGATTTATAGTTTTTAACTGAGTTAAAGTATT

AAACAAGTGCAGAGTCAAAAATTTAACAATTATTGTAAAGTTTATAATCAATA
AAGATAAAAAGAGTGGAGCAGAAATCCATCTAAAATGAATTCCGTTCCACTT
CATAGTTATTAAATATCATAACTTATATGCGTACATTTTCCAATTGATGT
CCTACTTCTCTGAAAATGCTACGCATTTGAAAGTTTATTTACCTAAGCGT
5 TGATTTTTACGATTTACAAAAGTTTCTTTGGCATCTTAAATAGAACGTTCT
AATTCAGGATATTACGACGTATTTTCATCTAATGTATCTTTCCAATACATT
ACTTCATCTTTGATGCTACTTATTTTGAAGGTTTACGTGTACGATTTTCCT
TCTTTGATGCTTTTGAATGATTGTTTAAAGAATGACGTGTAGATTTATCA
GCAAGAGCAGCCGCGCCACCAATTACGGCACCAATTTAAAATACCAGGTACT
10 AATTTATTTTCCATAGTTGAATTACCTCTCTTTTCAAATTTGCATCTTTTAC
GATGTAGTCTATTAAATTATCACAAGATGATTGCACCATCTCGAAAACACC
TTCAAATTTATTTGTGTAGTATGGATCTGGTACATCACTCTCTTCCATGTT
ACTAAATTTAGCAATTTGAACAATTTGTCCTTGTAAATTTGGATTGATTG
TTTGATATTGTCTACGTTACTTTGGTCCATAGCAATAATATAGTCAAAATC
15 ATCATCAGGTTGAAAAGTTCACTACCATACCATCATAAGGAATATGGTA
CTTCTGTAGAATTTTTTGTGTTCCGTTATGTGGAGGTTTCGCCTAAATTCCA
ACGTCCTGTTCTCTAGAAATGAACCTTTTATATCTGAAATACCTCTTTCTTG
TAGTCTTTGTCTCATGATAGCCTCAGCCATTGGAGAACGACATATATTACC
GAGACATACAAATGCTACATGTATCATTAGTATCACTCCAAACTATGTATT
20 TATATACAAATAGCTCATATTATTACAAAGTTCAAACCTTTTATACCTAAAA
ATTATATTGAAGATTTCAAAGACAGTTTAAACGAACCTGGATTAGCTAAAG
AAGGTAATTACGAATATTGTTTCTTTGAAAGTGAAGGTAAAGGGCAATTTA
AACCAGTAGGTGATGCAAGTCCCTTATATAGGGAAGTTAGATAGTATCGAAT
ATGTTGATGAAATAAACTTGAGTTTATGATAAA
25

Sequence 3455

step.1003d10.cons.ok

GTATCTAAGCCTAAAGCTAATGAAGCAGTAGTGACGAACGAGTCAACTAAA
30 CCAAAAACACAGAAGCACCACCTGTTAATGAGGAATCAATAGCTGAAACA
CCCAAAACCTCAACTACACAACAAGATTTCGACTGAGAAGAATAATCCATCT
TTAAAAGATAATTTAAATTCATCCTCAACGACATCTAAAGAAAGTAAACA
GACGAACATTTCTACTAAGCAAGCTCAAATGTCTACTAATAAATCAAATTTA
GACACAAATGACTCTCCAACCTCAAAGTGAGAAAACCTTCATCACAAGCAAAT
35 AACGACAGTACAGACAATCAGTCAGCACCTTCTAAACAATTAGATTCAAAA
CCATCAGAACAAAAAGTATATATAAAACAAAATTTAATGATGAACCTACTCAA
GATGTTGAACACACGACAACATAATTTAAAAACACCTTCTATTTCAACAGAT
AGTTCAGTCAATGATAAGCAAGATTACACACGAAGTGCTGTAGCTAGTTTA
GGTGTGATTCTAATGAAACAGAAGCAATTACAAATGCAGTTAGAGATAAT
40 TTAGATTTAAAAGCTGCATCTAGAGAACAAATCAATGAAGCAATCATTTGCT
GAAGCACTAAAAAAGACTTTTCTAACCCCTGATTATGGTGTGATACGCCA
TTAGCTCTGAACACATCTCAATCAAAAAATTCACCACATAAGAGTGCAAGT
CCACGCATGAATTTAATGAGTTTAGCTGCTGAGCCTAATAGTGGTAAAAAT
GTGAATGATAAAGTTAAAAATCACAACCCCTACGCTTTCACTTAATAAGAGT
45 AATAATCACGCTAATAACGTAATATGGCCAACAAGTAACGAACAATTTAAT
TTAAAAGCAAATTATGAATTAGATGACAGCATAAAAGAGGGAGATACTTT
ACTATTAAGTATGGTCAGTATATTAGACCGGGTGGTTTGAACCTTCCTGCA
ATAAAAACCTCAACTACGTAGTAAGGATGGCTCTATTGTAGCTAATGGTGTA
TATGATAAACTACAAATACGACGACTTATACATTTACTAATGTTGAT
50 CAATATCAAAATATTACAGGTAGTTTGTGATTTAATTGCGACGCCCTAAGAGG
GAAACAGCAATTAAGGATAATCAGAATTATCCTATGGAAGTGACGATTGCT
AACGAAGTAGTCAAAAAAGACTTCATTGTGGATTATGGTAATAAAAAGGAC
AATACAACCTACAGCAGCGGTAGCAAATGTGGATAATGTAAATAATAACAT
AACGAAGTTGTTTATCTAAACCAAAATAACCAAAATCCTAAATATGCTAAA
55 TATTTCTCAACAGTAAAAAATGGTAAATTTATACCAGGTGAAGTGAAAGTT
TACGAAGTGACGGATACCAATGCGATGGTAGATAGCTTCAATCCTGATTTA
AATAGTTCTAATGTAAAAAGATGTGACAAGTCAATTTACACCTAAAGTAAGT
GCAGATGGTACTAGAGTTGATATCAATTTTGCTAGAAGTATGGCAAAATGGT
AAAAAGTATATTGTAACCTCAAGCAGTGAGACCAACGGGAACCTGGAAATGTT

TATACCGAATATTGGTTAACAAGAGATGGTACTACCAATACAAATGATTTT
 TATCGTGGAAACGAAGTCTACAACGGTGACTTATCTCAATGGTTCTTCAACA
 GCACAGGGGGATAATCCTACATATAGTCTAGGTGACTATGTATGGTTAGAT
 5 AAAAAATAAAACGGTGTTCAGATGATGAGAAAGGTTTAGCAGGTGTT
 TATGTTACTCTTAAAGACAGTAACAATAGAGAATTACAACGTGTAACACT
 GATCAATCTGGACATTATCAATTTGATAATTTACAAAATGGAACGTACACA
 GTCGAGTTTGGCATTCTTGATAATTATACGCCATCTCCCGCAAATAATTCT
 ACAAATGATGCAATAGATTGAGATGGTGAACGTGATGGTACACGTAAAGTA
 GTTGTGTGCCAAAGGAACAATTATAATGCTGATAATATGACTGTAGATACT
 10 GGCTTTTATTTAACCTCTAAATACAATGTCCGAGATTATGTATGGGAAGAT
 ACAAATAAAGATGGTATCCAAGATGACAATGAAAAGGGAATTTCAAATGTC
 AAAGTGACGTTAAAAAATAAAAAATGGAGATACCATTGGGACAACGACAACA
 GATTCAAATGGTAAATATGAATTCACAGGTTTAGAGAACGGGGATTACACA
 ATAGAATTTGAGACGCCGGAAGGCTACACACCGACTAAACAAAACCTCGGGA
 15 AGTGACGAAGGTAAAGATTCAAATGGTACGAAAACAACAGTCACAGTCAAA
 AGTGCAGATAATAAAAAACAATAGACTCAGGTTTCTACAAGCCAATATATAAC
 TTAGGTGACTATGTATGGGAAGATACAAATAAAGATGGTATTCAAGACGAC
 AGTGAAAAAGGGATTCTGGTGTTAAAGTGACGTTAAAAAGATAAAAAATGGA
 AATGCCATTGGGACAACGACAACAGACGCAAGTGGTCATTATCAATTTAAA
 20 GGATTAGAAAATGGAAGCTACACAGTTGAGTTTGAGACACCATCAGGTTAT
 ACACCGACAAAAGCGAATTCAGGTCAAGATATAACTGTAGATTCCAACGGT
 ATACAACAACAGGTATCATTAACGGAGCTGATAATCTCACAATTGATAGT
 GGTTCCTACAAAACACCAAAATATAGTGTCCGAGATTATGTATGGGAAGAT
 ACAAATAAAGATGGTATCCAAGATGACAATGAAAAAGGAATTTCTGGTGTT
 25 AAAGTAACGTTAAAGGATGAAAAAGGAAATATAATTAGCACTACAACAAC
 GATGAAAAATGGGAAGTATCAATTTGATAATTTAGATAGTGGTAATTACATT
 ATTCATTTTGAGAAAACCGGAAGGCATGACTCAAACCTACAGCAAATTTCTGGA
 AATGATGATGAAAAAGATGCTGATGGGGAAGATGTTCTGTGAACGATTACT
 GATCATGATGACTTTAGTATAGATAATGGTTATTTTGACGATGATTACAGAC
 30 AGTGACTCAGACGCGATAGTATTCAGACTCCGACAGTGACTCGGACGCA
 GACAGCGATTCTGACGCAGAC

Sequence 3456

35 step.1003d11.cons.ok
 GCATTTGCTAAGGGACTTGTTCATCGTTTGGGCCATTTTCCTTCAGCAACA
 ATACAACATATAGGCTATGGCGCCGGCAGTAAGGATTTTGATTTCCTTAAT
 ATATTAAGGGTTATTCAATTTGAATCTGAATTCGAGCAACAAGATAGCGTC
 CAAGTAATAGAGTGTCAAATAGATGATATGACACCTGAAGCATTAGGTTAT
 40 TTTATGAATAATGCGTTAGAGCAAGGTGCTTTAGATGCTTACTATACGCCT
 ATATTTATGAAAAAAGTCGCCCAAGCACGCAGTTAACGTTAATATGTAAA
 TTACATGATAAGACATATTTTGAACAACCTTATCTTACAAGAAACAAGTTCT
 TTAGGCGTCAGAAGTACTTCTGTTAATAGAAAGACCTTGAACCGCGCATTC
 AAAATTCTTTCTACACAACACGGCACTGTTTCCATTAAATTTGGCCTACAA
 45 AATGGAAAAATTATGAAAATGAAACCCGAGTATGAAGATTTGAAGAAAATA
 GCTAAACCTACAAAACAACCGTTTCAAGTAATTCATAACGAGGTATTACAA
 CAACTCTATCAAACATATCATATAGGAAATATACCTCAATAAATTAGTTTT
 ATTCATTTAAATTTTGACATCAAAAAAGGATAGGGACAGTCATTCTACAT
 AAAACAGTCACCTTATCCTTTTATCTATTCAAATGCTCGATATAGTTCTTC
 50 ACTGAATTTATTTCCACATTTCAACGTTATGATTTTGTAGAAATTCCTTTAA
 AATTTGTTTGTATCATCACCGATTTGGTCGATGATAAATGATGTTTCAT
 AGACTTATCCATCATATTAACATGTTCTGGCATGCTCTTATAACCACGTCT
 AATACTTTTGTAACTAATAAATAGCATGCAGTCACACCCGCATAATATGG
 CCCTTCTTCACTACGTTCCGTCGTCACCCAACATAGCCAGTATGCTTTAGC
 55 CTCCTCACCTTCAACTTGTGATTTGTGATTAATCCACTTGACACCCCTTTC
 AACTTCAGCACGTCGCTGCATACCACCAATATCAACAAATGCCTTCGATT
 TTCAACATCTATAAAAAACCGGGCAATATTATCTAAACTAATCGAACCAAT
 ATTGGTTCCTTTATGACCATCTAGTGGATCATTTTATATAATATTGAAAT
 GAAACCTTTTTTCTCAGACAAACGAGATACCTCCTAATAAGTCATCAATA

AAGACTATATATCTTATTCTTAACTGTTAATTTAATTCATTAAGTAAGTTA
 ATAATAGAACCTTTAATTTGATAATCTTTAATTTTAAATCATCTCTTTA
 GGATAATATAATTTATGGTCTTCTCTATTAATCCCTGTAATACTATTAATC
 ACCAATGATTGATTACTTATTTCTGTAATTCACCATTACTTTTTAATAAA
 5 TGAATCGGTTTGGGATTTGAGCCTGGGCGATCATAATCATAAGGTAAATCT
 GAAAATGCTTCACTTACAAAGTAATAATCAGGATTAATACCACCCGCTTCA
 AATAATCTTGCAATTCGGAATGGTAATAATGAACCGTCGAAAGGAATA
 TATTTAAATAAATCTCGATTGATAAACCGTCTTGATAAATCACTTAATATT
 GTATCATTTTCATGAATCCAATTTCTTCAAGTAATACAATACAACCTGCTTCA
 10 TCAAGTCTACATATTGTTCAATCGTCATTGTTCCCTTCAAAGAATGGTATA
 AAGTCTTTTGGATACATTTTAAATTCATATCCTTCATTATAAAGCTGCTTA
 GCTCGTTTTAAACAATTGTTTAAATAAACTTCCCCACCACGGCTTACTGGA
 TGAAAATATATTTGCCAATACATTTGATAACGACTCATAATAAAATTTTCG
 ACAGCATGCATACCCTATCTTTAATTAACACTTCTTCTTTAGATGGTCTC
 15 ATCAACCTTAAAAATACGCTCCATATCAAATGAGCCATACGTTACGCTGTA
 AAATATGCATCTCTTTGTAAATAGTCCATTTCGATCAGCATCAATTTGTGAG
 GAAATCATAGAAATAACAAGCTTATTTATCGTGCCTTTTATTAATTACATCA
 GCTACTGTTTGGAAAGTTATCAGAGACTCGACTTAGTACACCATTAAACC
 TCAGTGTCTCCTGTGATAATAGCTTGTGTAAATGCTTCATGGTCTGTATTA
 20 AATATTTTTTCAAACCTATGAGAGAAAGGACCATGACCTAAATCATGTAAC
 AACGCAGCACACAACGCAAGTGGCCGATCAGTATTATCCCACGCGTCTTGT
 CCTATAAATGTCTCATCAATCATTCGGCGCACAAATTCATAAACACCCAAA
 GAGTGTCCAAAACGACTATGTTTCGGCTGTATGAAATGATAAATACAAAGTA
 CCTAGTTGTTTAATACGTCTCAAGCGTTGGAATTCCTTTGTTTAAATTAAG
 25 TCCCAAATCAATTGATCCTGTACGTGGATAAATCGATGAATTGGATCTTTA
 AATACCTTTTCTTCTGTAAATTTACTCGTTGCATATGATGATAGCGTCATA
 CTTTCCCTCCTATCCTATATATTTCCAAAAACATTTATTTTGTTCCTCA
 TATTACTAGACGAAACAATGTTAATTATTCATATGTTTGTAAATGATTT
 CTTTCAATTAAGCTAAATCCATGATTTTACCATAACATTTTCAGGTTGATATGA
 30 CGGAATAATCTCGAATCCATGATTTTGTATAATATTCGATGCCATTACTATT
 ATTGTTATCAACTTCTAAATAAACTGTTTCATACCTGGTCTTTAAACCTTTT
 CAACCCCGCTTCTAGCAAACGAGTTCATACCTCTATGTTGAGATTTCAGG
 TCTAACATAGTGTGCAGATAAATAAAGCTCTTACCATAAATGAAATTTGC
 AAAACCAACAATCTCACTATCTTCTTCTGCCACTAAAAATAATTGCTCATT
 35 TAGTCTCTTTTAAATGCGTTTCATTATATGAAGCGGCCAACAGTTCATT
 AACTGTTGAAGCAGCATAAATATTTAAATATGTTTATAACCAAGCTTTTGT
 CGCAACATCTCTTATCGCTACAACATCTTTTTCAGTTGCTTGTCTGACACT
 ATACATGTGACTTCTCCCTAACTATTTACTTCACTATTAATCGATGGGTG
 GATTGGTAACCTGTGCAAATAATGTTTAAATAAATAAATAGTGACTCAG
 40 GCTTATCATTAAGAAGTGCAATAGGTGCATCATATCCCATCGCTCCGATGG
 GATCTTTTTTATTGATGACTAAATCTGTCATATATTTATTTCTTTAGAGCAC
 AGTGGCGATGATATC

45 Sequence 3457

step.1003d12.cons.ok

ATTTTTCAATCATCGATAAACTTAAACTAATTCATTCTCATATTGGGAGAT
 CTTAATATTTTAAATTCATATATAAATAGTAATTTTAGACACTGATCAATCT
 TTTAAGTTGGTTATTTTAAATATCAGGTTCCAATTTTTTCTTTTACAACCTT
 50 AACAAATTCACCTTACCTCATATACCTTCTAAAACAGGTTTCATAAATTAATG
 ATAAATGACGTTTGTATCACATCAAAGGCATTTGCCTATATTTTCAAAAA
 ATGTGTCTTTTTTTTCAATCTATTTTCGATATATATAGTGAAAGGAGGTTT
 TATTATGAATGAAGTCAATCAAGTGACGATTATCTTACTCAAATCAAACA
 TAATTCCTGAAGGTAAAGTAGTCGAACATAAGCGTCGTTTTTCTAACATTAA
 55 TCCGCTTGTTTCAATGATCAAAATCAAATCTTTTCGTTCCATCATTGAACG
 TATCAGTGCGGAAGAGTATGACAAAGTCGAGATCGTTAAACAGAACTCTTT
 AATTTAAGGAGGATACATCAGTATGTCAAAAACCTTGGAACCTCGTATTTAA
 ATCAAATCTTAAATAAACCCGTCAAATTAATTAACCCGATTTTAAACGCTAT
 TACTACAGAACAATTAATTAAGAAAGTATGAACCAACTTTTAGAATTAGA

CATTTTAAAGATTAGTTTGGGAAAGCCTGTCAAAATTTATGCTGCACAAC
TATTGATAAAAGTACAACGTATATTTGAAGATAAAAAATTAATTCATAAA
ACATTATTTTGTGACGCTTTATTTAAAGTACTAAAATCATAAAAAAGACTTA
TCGAAGAACAAATTTTTTAAATTTGTTCAAGTCTAAAAAGAAGCTTAATTT
5 ACATTGTTCTTTTATTAATATGAAGTGAGATATCAAATTCAAAATGATGTC
TCGCTTTTAAATTTCTAAAACCTGTTTAAATGTGCATTTTATTATAATGAT
TCATCAAAATAGTATCGGCTGTATGACACCTCGGAACATTCAACAACCTATT
TCTATACCTTGTCTATGGCTCAATCATTTTAGAAACATGATTATGAATGGT
ATGATAATCATATTAATTTGAGAAAGTTGGTTTTATCATTGTTTTTCTA
10 TATCTTTTAGGAATTTTGTGTTGGCATGTTACTTCTTATTCAAACCTCAATC
AACTCAAGGCTAAGCCAATTTACCCGTTTCATCTTTTATGCTTCAACCATA
TCTTTTGTGTCGGAACAATATGCTTACTTGTGCTTAATATCATTATTCAT
CCACAAGTATTAACACCAGAAATCTTTTCTAAGCAAACACTTAACATACT
TGGGTCTTAGGAGGACTATTAGGTGTTATCTATTTAACTGGAACTTATTA
15 TTATTACCAAGATTAGGCGCAGCACTTACTGTTGTTATTACGGTTACAGGG
CAAATTTATTGGGTGTAATCATTGATACATTTGGATTATTGGGCGCCCAT
CAACAATCTTTTACATTTATTTAAAGGTGTTGGAATTATTTTTTAATTA
GGAATTTATTTATGAATTTATGTTTCGTAGACATCCTGTTAATAGACATAAG
AATACACCAATAGTATTTCTGGTTGCTTATTTGGATTTCGTATTTGGTTTCGCT
20 CCGCCAATCCAAACAACAATCAACAGCACACTGGCTCAACACACACACTCA
TCCATATTTGCTTCACTCATATCATTTAGCGTTGGTACAATAGCGTTATTT
ATATTAACATTAGTATTTAATCGTAGTCTGAAAATATCATCTACTCATAAA
ACACTAGGTAAGATTAAATCTATTTATTTTATTGGAGGGATAC TAGGAATG
GCATTTCGTAACATCTAATATTTATTTAATGCCCTTCTTAGGAGCAGCTTTA
25 ACTACTATAATTGCTATGATGGGCCAAATGATTTATGGGTATTATTATCGAT
CATTTCCGTTTGTGTTGGGTTCCACTAAAACAAAATAACATCACGTAAGTTA
GGTGGATTATTATGTATAGCTATAGGTATTATTTTATTACGATTATTTTAA
ATTCTTATTCTGAATAAAAAATTATCCGCTTATATTAATGTAGATTCCCTCA
TACAAATAAATGTCTATTTTCGAACTGCCAACAAGATTTTGTCTTTGTGTTGG
30 CAGTTTTTATGAGTGATTCTTCTTATAGTAACATTTATGTAGAATCATATA
TTTTTAAATTAAGATAGCCTTCTATTTTTTCAATTTATTAAATTGATTAC
TTTATTTTTTGAATAAAACAATAGTAATTGTGCACCATATTGTTCCGTCT
CTTCTCTGTAGTAGCTAATAACGTATCGTACCTTGCTTTAATGCGTTTG
ATTGTTTTAATTCATCATCGCGATGACCATCGCCATAAGACCACCTTTCA
35 TGTCAGTGGTACCTCTGCCGAATAACTTGCCATCTTTATCTGTGAGTTCAA
ATGGAGGAAAAGTCCAATCATCATGGTCACCAGCATCTACAACATCCATGT
GTCCACTGATTGCTAAGACTGGAGCGCCACTACCAATTTAGCAACTAAAT
TAGCGCGTGAATCATTAACCTTAAACAATCTTAGAATCAATATCATATTGAC
TTAACAGGCTTTTTAAGTACTCACATACTTTAATTTTCATGATCATTTCTCAG
40 TTTGAATCTTGACGATATCAGCTAATAATCTTATTTTGTCTTGCTCACTCA
ATACAGTCATTTACAGTCCCTCCTCATTTTTTAGTAATGATTTTCACTTTTT
AATAAAAAATAAAAAATCCTTGTACACTTAATATTATCAAAAACTTTTATTA
ATTAAAAATTATATGCTAAAAATAAATGGATTTTAAATAAAAAATAAATATT
AACTTAACATTTCTTATAAAAAACAACAATGTATAGAAATATTAAAAATTT
45 TTCAATTATTTATAGTTTGAATCGTGTTATAAAATGAAAAGCTATGAGATTA
AGATGTAATAATACACACATATATAAATATAACTTCGTGCCACGCAGACGG
TTAAAAATCATGTATAATAGCTAATAATACATATTAGGAGATACATAATTAA
TGATGAAAAATAAATTAACATTAAGAGAAATCTATTTATCGGCTCAATGC
TGTTTGGTCTTTTTTTTGGTGCTGGAAATCTCATTTTTTCCAATTCACCTAG
50 GTCAAACCTGGGGGGCAAAATGTATGGACCGCAATTTAGGATTTCTTATCA
CGGCTATCGGACTACCTTTTATAGGAATTATAGCGATAGGTGTATCTAAAA
CAAACGGGGTCTTTGAAATTTCTCAAGGATAAGTAAAAATATATGGTTATT
TGTTTCAAAATGGCTTGTATCTTGTATAGGTCCGTTTTTTTGGCTTGCCAA
GACTTGGACGACGTCATTTGAAATAGCATTTTACCATTATTTTCATCTG
55 GTACGGCCCAAGCGTTGTTGCCTATTTTTTAGTATTTTATTCCTCGGAGTAG
CGTGGTTATTTTTCGCGTAAACCTTCTAAAAATATTAGACTATATTGGAAAT
TCCTTTAGAGCACAGTGGCGATGATAT

Sequence 3458

step.1003e02.cons.ok

GTCTCGCAGTCAAGCTCCCTTATGCCTTTACACTCTATGAATGATTTCCAA
CCATTCTGAGGGAACCTTTGAGCGCCTCCGTTACCTTTTAGGAGGCGACCG
5 CCCCAGTCAAACCTGCCCGCCTGACACTGTCTCCACCACGATAAGTGGTGC
GGGTTAGAAAGCCAACACAGCTAGGGTAGTATCCACCAACGCCTCCACGT
AAGCTAGCCCTCACGTTTCAAAGGCTCCTACCTATCCTGTACAAGCTGTGC
CGAATTTCATATCAGGCTACAGTAAAGCTCCACGGGGTCTTTCCGTCCTG
TCGCGGGTAACCTGCATCTTCACAGGTAATGATTTACCGAGTCTCTCG
10 TTGAGACAGTGCCCAAATCGTTACGCTTTTCGTGCGGGTCGGAACCTACCC
GACAAGGAATTTTCGCTACCTTAGGACCGTTATAGTTACGGCCGCCGTTTAC
TGGGGCTTCGATTTCGTAGCTTCGCAGAAGCTAACCACTCCTCTTAACCTTC
CAGCACCGGGCAGGCGTCAGCCCCATATACATCACCTTACGGTTTAGCAGAG
ACCTGTGTTTTTGATAAACAGTCGCTTGGGCCATTTCACTGCGGCTCTTCT
15 GGGCGTGAACCTAAAGAGCACCCCTTCTCCCGAAGTTACGGGGTCATTTT
CGCGAGTTCCTTAACGAGAGTTTCGCTCGCTCACCTTAGAATTCTCATCTTG
ACTACCTGTGTCGGTTTTCGGGTACGGGCACCTGTTATCTATCTAGAGGCTT
TTCTCGGCAGTGTGAAATCAACGACTCGAGGAAACAATTTCTCTCCCCAT
CACAGCTCAGCCTTATGAGTGCCGATTTGCCTAACACTCAGCCTTACTGC
20 TTGGACGTGCACTCCAACAGCACGCTTCGCCTATCTACTGCGTCCCCCA
TCGATTAAACGATACTAGGTGGTACAGGAATATCAACCTGTTATCCATCG
CCTACGCCCTGTGGCCCTCAGCTTAGGACCCGACTAACCCAGAGCGGACGAG
CCTTCTCTGGAACCTTAGTCAATCGGTGGACGGGATTCTCACCCGTCCTT
TCGCTACTCACACCGGCATTCTCACTTCTAAGCGCTCCACATGTCTTTCG
25 ATCATGCTTCGACGCCCTTAGAACGCTCTCTACCATTGTCCAAAGGACAA
TCCACAGCTTCGGTAATATGTTTAGCCCCGGTACATTTTCGGCGCAGTGTC
ACTCGACTAGTGAGCTATTACGCACTCTTTAAATGATGGCTGCTTCTAAGC
CAACATCCTAGTTGTCTGGGCAACGCCACATCCTTTTCCACTTAACATATA
TTTTGGGACCTTAGCTGGTGGTCTGGGCTGTTTCCCTTTCGAACACGGACC
30 TTATACCCATGTTCTGACTCCCAAGTTAAATTAATTGGCATTCGGAGTTT
GTCTGAATTCGGTAACCCGAGAGGGGCCCTCGTCCAAACAGTGCTTACC
TCCAATAATCATCACTTGAGGCTAGCCCTAAAGCTATTTTCGGAGAGAACCA
GCTATCTCCAAGTTCGATTGGAATTTCTCCGCTACCCTCAGTTCAATCCGCT
CACTTTTCAACGTAAGTCGGTTCGGTCTCCATTTCAGTGTTACCTGAACTT
35 CAACCTGACCAAGGGTAGATCACCTGGTTTCGGGTCTACGACCAAATACTC
AACGCCATTTACAGACTCGCTTTCGCTGCGGCTCCGCATTTGCTGCTTAAC
CTTGCACTCAGATCGTAACTCGCCGGTTCATTCTACAAAAGGCACGCCATCA
CCCATTAACGGGCTCTGACTACTTGTAAAGCACACGGTTTCAAGTTCTCTTT
CACTCCCCTTCCGGGTACTTTTCACCTTTCCCTCACGGTACTGGTCACT
40 ATCGGTCACTAGAGAGTATTTAGCCTTAGGAGATGGTCTCCACAGATTCCG
ACGGAATTTACGTCGCTCCGTCGTACTCAGGATCCACTCAAGAGAGAATAT
GTTTTGCACTACAGGATTTATACCTTCTTTGATTCATCTTTCCAGATGATT
CGTCTAACATGTTCTTTTGTAACTCCGTATAGAGTGTCTACAACCCCAAC
AAGCAAGCTTGTGGTTTGGGCTCTTCCCGTTTCGCTCGCCGCTACTCAGG
45 GAATCGATTTTTCTTTCTCTCTCCCGGTACTAAGATGTTTCAGTTCTCC
GGGTCTGCCTTCTGACATGCTATAAATTCACATATCGATAACATGACATAA
CTCATGCTGGGTTCCCCCATTCGGAAATCTCTGGATCAACGCTTACTTACA
GCTACCCAAAGCATATCGTCGTTAGTAACGTCCTTCATCGGCTTCTAGTGC
CAAGGCATCCACCGTGCGCCCTTAATAACTTAATCTATGTTTCCACCATAT
50 TTTGAATTGTTATTTCAAAATAAATAGCTAAACTAGTTATTAATCTTGTGA
GTGTTCTTTTCGAACACTAGCGATTATTTATGAATTCAAGCTTATTTAAAC
TCTATTCACTCGGTTTTGCTTGGTAAATCTTACTTACTTATCTAGTTTTTC
AATGTACAAATGAATGTTAATAAACATTCAAACTGAATACAAATATGTCAC
GTTATTTCCCTCATCTTCGTAGAAGATGTTCCGAATATATCCTTAGAAAGGA
55 GATGACAAGTTTCATATAATCATTTGTCTAAATTTCTGATATTTGGAATTTTG
TCATGATTTAAAGTCTCAAGCAGTTTAGCTTTTTTAAATTTTGAACAGTA
TATTCACCTTGGAAAGCTACATCATAGTGTGTGCCTCCATTTTGAATTTTC
GCTTCCATTGCTTTCATTTGAATCAAAGGTTTCATATACGACTTGTATACCT
GTTTCTTTTTTAACTTCTTAATTAAGCTAGGATCAATATATTTCTCCCCAA

TTGTAACATATAAATTCTCATTGAGTGGGTATTATCTTTTGATTAAAC
 CAATGACTAATCATTAACATAAAATACCAATAACTATAGCAGAAATGATA
 AGTTGAAAAAATCGCTTCATTTCATTAACACCTCGCTTCATCTTTTCTGAC
 GATTAACGATATGTTGAATAAAGTAATATCCAAATACTCCAAACATAATAG
 5 CAATAAATATTATTGTAGAAATGGCATTAACTCTCCATACTTATACCTTTTC
 GAGCCATAGCATAAACTTCTACAGACAATACACTAAATCCGTTTCCAGTAA
 CAAAGAAGCTTACCGTAAATCATCTAGAGAATATGTTAATGCCATAAAGA
 ATCCTCCTATAATCGAGGGCATAATATTAGGTATAATAATGCTAGTTAATA
 ATTGTGATTTCATTAGCACCTAAATCTCTTGCTGCATTAAACATGTTATCAT
 10 TCATTTTCATATAATTGGGGTAAGACGATGATAACAACAATTGGAATACAGA
 ACGCTATATGAGATGCTAGCACTGTTGAAAATCCTAAACCTAAACCTGTGA
 AATGTCCATATGCAGTGAACATAATTAAAAAAGAAGCACCATGACTACGT
 CTGAAGATACCATTAAATACATTATTCATCGTTAATAGCGTAACCTTAAACC
 GTTTGTTGCGTAAGTAATATAAAGCGATCGCACCATAATGTACCTATAACTG
 15 TAGCAATCGATGCAGATAAAAGTGCCACTGCAACTGTATTAATAATCACC
 ACATCAACCGATCATTATGAAAAAGGGAGTGGTAAATGTTCTAAAGTAAAT
 GTTCGAAGTACTATTATACCCGAGAAATTAAGAATAAATCATTAATAA
 AGAAAATGGGAATATAAAGCACACTTATTAATAGAGCTATATATAATTTTC
 CGTACCATTTACACTAATCACCCTTCCCATTAGATGATTTTGATTGTTGT
 20 AATAATGAGTAAAAAGGCCATAAAAAATAATGAGAAAAAGTGCTATAGTGGA
 ACCTAATCCATAATTTTG

Sequence 3459

25 step.1003e03.cons.ok
 TTTTGTAAATTTACTGCTTCAATCATTTTCACAAATAAGTTGTTTAAATTC
 ATTGATATTAAGGTGTAATTGAGTGTCACTTTAGCGTGGTTAATAGTTAA
 TAAAGCTTTTTCAAAAGATTTTAATGCGCCTTTTAAATCCCACGACGATA
 GTGATAATTAGCTGTTGCAACAAAATGAGACTTACTACTGCATCATGTTT
 30 TGAAAAAGTATTTTGTGATTTTCCAAGCATCTTCTAAAATATCATGACATAG
 GAAGTAATGTTGATGCTTATGAAATTGATAATAATAATCCATTAAATGATTG
 ATCCATTAGATTTTACCCATAATTTAAAAGTTGTATGTATTTCATGCTATAA
 TAATTTAGTGAATATGCATTACGATGCAATTTGATATAGAGGTAGATGT
 AATGTATGAAGTAAACTCGATGCATTTAATGGTCCATTAGACTTATTTATT
 35 GCATCTAATTCAAAAATATGAAATTGATATTTATGATATCCCTATGAAAGC
 CTTAACTGAACAGTACATGCAATATGTTTCATGCGATGAATCAGCTAGAAAT
 TAATGTTGCTAGTGAATATTTAGTTATGGCATCAGAATTACTAATGATTAA
 AAGTAAATTCACGATGAATATAACCAAAAATCCAACACTAAAAACATGCT
 GTGCGGCTTCTGCACTATGAACGTGTTGTTTACCATTAACTGTTGTAGTAA
 40 AATTAGCATCCATCGCTCCAGTAATTGAAAATACCAAGAAAACATAAAACA
 ATAGACTAGAAAAACTTTGTAAAATGATTCCCTATCAATGATGTTATATTA
 TTGCTTTATTCATCATAATTCCTCTTCATTTTCTACGTATATCCTTACAA
 TCAGGTTTATGTTCTTTAATAATAATCTTTAGGATTTATATCTCGTTTATAG
 TAGACTTTCTTTCCCTTTATTTTCATATGCAACATAAAATAGTTGCTCTTTA
 45 TTTTGAATAATTTATAACTAATAATAATCAATTTTCCATATAAAGAAGTA
 GATTTTTCCGCTTCACCAGAAATAATTAAAGGTGTTCTTGCTTCGTCGATC
 AATATGGAATCGACCTCATCAATAATTGCAAAATGTAGAGGACGCATTACT
 CTCTCTCAGCATAGTTACCATATTATCTCTAAGATAATCAAACCCAAGT
 TCATTATTCGTACTATAAGTGATATCTTGTCGTAAGCCTCACGTTTCTCT
 50 TCAGTTGACTTACTATTTAAGTTCAAACCTACAGTTAAGCCAAGATAGTTA
 TATAGTTACGACATTTCTTCACTTTGTGAACTTGATAGATATTCATTGACT
 GTAATAACATGTACACCTCTACCAGCTAAAGCATTCAAATACGTCGGCATG
 GTTGCAGTCAATGTTTTCCCTTCACCTGTTCTCATTTCTGCAATATCACCT
 TTATGTATAGCAATACCACCATTAATTGTACTTTATAAGGAATCATATTA
 55 AATACTCTCTTGACCCCTCACGTACAAGTGCAAGCTTCAGGTAAAAAT
 TTATCTAAATAATCATTTTGTGTTTTTACATCTTGATGATTTAGTGATCAC
 CACAATCCTTTTATCATGTAACCTAGTAAATTTAAGTTTCAACTTTT
 ACTTATAGTCTTTGACTTTAATTATAAATCAAAATTCGTCATAATAAGACG
 TGCTAAATTTACAATAGCTAAAACATGATAGAGGTGACTTTTATGTCTACG

AATAATGAGATTGAATTTAAACAAATACTAGATCAAGATACTTACTCAAAA
ATCTATGAACACTATTTCAAAAATCAATCACCTTTTAAAGCAAACATAATTTC
TATATCGACACAGAGAATTTTAAATTAAAAACAGCATCATGCTGCTTTGCGT
ATAAGGGTAAAAGATTATATGTTTGAAATGACTTTAAAAGTCCCTGCTGAA
5 GTTGGATTGACAGAATATAATCACTCAGTAAATATAGAACCTGAACCTGAT
ATGTCACCTTCAACTTTCTCAATTACCCAACGATATTAGAAATATTTTAGAA
CAGGACTTTAATATTTTAGAAAAATGAGCTTAAAGTACTAGGAAACTTAACT
ACCTATCGTTTAGAAACCGATTATCAAAAATGAATTACTAGTATTAGATAAG
AGTGAATATCTCGGCAAACTGATTATGAATTAGAGTTTGAAGTTCATTCT
10 TATGATGAAGGATATTTCAAAATTTAAAACTTTACTTCAACATTTTAACTCT
CAACATCAAAAACCCCTTGAATAAAGTGCAACGTTTFTTTTCAAGAAAAACAA
AATGCAAGTGATAAAGAGTAACAGTTGTGCGATAAGGCAAAATTAATTTTAAA
TGTA AAAAATTCATGTTATATTAGATATATATTAATGGAACACGGTGATATT
ATGTCTAAAACACCATATGAGTTGATTGGTCAAAAAGCCTTGTATCAAAATG
15 ATTGATCATTTCTATCAACTTGTGCGAGAAAGATTCTCGTATCAATCATTTA
TTCCCAGGCGATTTCAAGGAAACAGTCGAAAGCAGAAGCAATTTTGTACA
CAGTTTCTTGGAGGTCCTGACTTATATACCCAAGAACATGGTCATCCCATG
TTAAAACGAAGACATATGGAATTTACAATTAGCGAGTATGAACGTGATGCA
TGGCTTGAGAACATGCATACTGCTATTCAACACGCCAACTTCTGCGGGT
20 GTAGGCGATTACTTGTGTTGAGCGATTAAGACTTACTGCAAAATCACATGGTA
AATTCCATAAATTTAATTGTAGGTGAAAAATACATGGCTGAAGAATTAAGAAT
CATGGAGAATAAGAGTCGTGAAGATACTAATCTATCACCTGTTAGCAAAAT
AGAAATCTATCTTTTTTTGATCCTTTTAGCAAAGATTGTTTTAAATTATC
TGCAATCTTTATCAAAAATTAAGAATTGAATATAATAAATATATAAAGGTAAG
25 ACATATTTTAAACCCCTTCTTTAAAGGTATTAACTAAGTGTCAAGCTCAAAG
TACTTCAGATTTTGACAATATTGCACCTGCCTATAAAGCCGCTGAACCTCA
AGGTGCTATCAGAGCAGAAAGATTTATACATTTAATGCAAAATGAAATCAT
TCCAAAACGTGATATTATTACCGAAGATATGATTTCTGATTGTATTAATAA
TGCCGGCATTTGACTATCAAGTTTTTAAAGAAGACTTGCAAAAGGACAAGTT
30 GACTGACAGCTTGAAAGTTGATCTTCACATTGC

Sequence 3460

step.1003e04.cons.ok

35 TATATCGGCACTGAAAAGCCTATGCTAAAAAATAAAAAATATCGCGCTTCTT
TTTGAAAAGATTCCACTAGAACACGTTGCGCATTCGAAGTTGCCGCACAT
GATCAAGGTGCACACGTCACTTATCTTGGACCTACAGGTTCTCAAATGGGT
AAAAAAGAACTGCTAAAGATACAGCACGTGTACTTGGTGGTATGTATGAT
GGTATTGAGTACCGAGGTTTCTCTCAACGTACTGTAGAAACATTAGCGCAA
40 TATTCAGGTGTTCCGGTATGGAATGGATTAACCGATGAAGATCACCCATACA
CAAGTGCTTGCTGACTTTTAACTGCTAAAGAAGTATTGAAAAAAGAGTAT
GCTGATATCAACTTTACTTATGTTGGCGATGGACGTAACAATGTTGCTAAC
GCATTAATGCAAGGTGCTGCCATTATGGGTATGAATTTCCATCTTGTTTGT
CCTAAAGAACTCAATCCGACAGAAGAATTATTAATCGTTGCGAACGTATT
45 GCGACGGAAAATGGCGGTAACATTTTAATAACAGATGATATTGATAAAGGC
GTGAAAGATTCTGATGTTATTTATACAGATGTTTGGGTATCAATGGGCGAA
CCTGATGAAGTATGGCAAGAACGCCCTTAACTTTTAAAACCATATCAAGTT
AACCAAGCATTTATAGAAAAAACAGGCAATCCAAATGTTATTTTGAACAT
TGTTTTACCTTCTTTCCACAATGCAGAACTAAAATTGGTCAACAAATTTAT
50 GAAAAATATGGCATTTAGTGAAATGGAAGTCACTGATGATGCTTTCGAAAGC
AAAGCTTCTGTAGTATTCCGAAGCTGAGAATAGAATGCATACAAATTA
GCGGTCTAGGTAGCAACTTTAGGAGAATTCTAAATTATAGAAGGAAGTGAG
TGAAATGGCTAAAATTGTAGTAGCTTTAGGTGGAACGCTTTAGGAAAATC
ACCACAAGAACAACTTGAATTAGTAAAAAATACAGCTAAATCCCTAGTAGG
55 ATTAATTACTAAAGGTCACGAAATTGTGATTAGTCACGGTAATGGACCACA
AGTAGGAAGTATTAACCTTGGTCTGAATTATGCAGCTGAACACGATCAAGG
TCCTGCTTTTCCATTGTGTAATGTGGCGCTATGAGTCAAGCCTACATCGG
CTATCAACTTCAAGAAAGTTTACAAAATGAACCTCATTCATGAGGCGATAGA
TAAGCAAGTTGTCACACTAGTTACCCAAGTAGAAGTTGATGAAGGCGATCC

AGCTTTTAAATAGTCCAAGTAAACCCATCGGTCGTCTACACTAAAGAAGA
 AGCAAATCGTATTCAACAGGAAAAAGGTTATCAATTTGTAGAAGATGCTGG
 TCGAGGTTACCGTCGCGTTGTACCATCACCACAACCAATATCTATTATCGA
 ACTGGAAGTATTTAAACTCTAGTAGAAAATGACACACTCGTCATCGCTGC
 5 AGGTGGAGGTGGTATACCAGTCATTCGCGAACAGCATGATAGCTTTAAAGG
 TATAGATGCCGTCATCGATAAAGACAAAACAAGTGCATTATTAGGTGCTGA
 TATTCACGTGTGATCAACTCATTATTTTAAACAGCGATTGATTATGTTTATAT
 CAACTATCATACTGACCAACAACAAGCACTTAAACAACAATATAGATAC
 GCTTAAACAATATATTGAAGAAGAACAATTTGCCAAAGGCAGCATGCTACC
 10 TAAATCGAATCTGCCATCTCCTTTATTGAAAATAATCCTAACGGTAGCGT
 GCTCATCACATCATTAAATCAATTAGATGCAGCACTAGAAGGTAAATTTGG
 CACACTCATTACAAAGTAACGTTTAAACGTATAAGTAGAAATTAATTTAAT
 CATATTATACAGTGCATCCTGTTATTCCACTCCTCTCCATTTATACAGG
 GTGCACTTCTTAATCAATCATGTAAAATACTAACCAAAATCAATTGATAG
 15 GCTTTTATTTAGGTTAACTGTGAAGGAGATGAAACCGTTGGAACAAGCG
 ATCAATGATAAATAAAGAAAAACGTTTAACTTTAGAATGCCAGGTGCA
 TTTATGATTCTCTTTATCCTAACAGTTGTGCGAGTTATAGCAACTTGGATA
 ATCCCCGCGGGTGCATACTCAAACTTTTCATATGAACCTTCATCCAAGAA
 TTAATAATTTGCAATCCTCATCATCAAGTAAAAAAGTTTCTGGAACACAA
 20 AAGGAGCTTGATCGATTAGGAGTTAAATCAAAATAGAACAATTTAAATCT
 GGTGCAATTAATAAACCCGTTTCAATTTCTAATACTTACGAACGCTTAAAA
 CAACATCCAGCTGGTCTTGATCAAATTAAGTAGCATGGTTAAAGGAACC
 ATCGAAGCCGTCGATATTATGGTCTTTATACTTGTCTAGGTGGACTGATT
 GGTATTGTTCAAGCGAGCGGTCCTTTGAATCAGGATTGTTAGCACTTACT
 25 CAAAAACGAAAGGCCACGAATTTATGTTGATTATGTTTCGTAGCAATTTTA
 ATGATTCTGGGTGGAACACTATGTGGCATTGAAGAGGAAGCTGTAGCGTTC
 TATCCTGTACTCGTTCCAATATTTATTGCGCTTGGATATGATTCTATTGTC
 TCAGTCGGTGCAATTTCTTAGCAAGCTCTGTGGGTAGTACATTCTCAACA
 ATCAACCCATTTCTCAGTCGTCATTGCTTCTAATGCAGCAGGAACAACTTT
 30 ACTGATGGTCTTTATTGGAGAATAGGCGCTTGTATCATCGGTGCCATATTT
 GTTATTAGTTATTTATTCTGGTATTGTAAAAAATTAATAAGATCTTAAA
 TCCTCTTATCTTATGAAGACAAAGCAGCATTTGAAAAACAGTGGTCTGTG
 CTCCATGATGACGGTTCTTCTGAGTTTACATTACGTAAAAAGATTAATCTT
 ACGCTTTTCGTCCTACCATTCCCTATTATGGTTTGGGGCGTCATGACACAA
 35 GGATGGTGGTTCCAGTCATGGCATCTGCATTCTTGATCTTTACCATTTGTC
 ATCATGTTTATTGCTGGAACAGGACAATATGGTTTAGGCGAAAAAGGCACT
 GTAGATGCATTTCGTTAATGGCGCTTCAAGTTTAGTAGGTGTATCTTTAATC
 ATTGGTTTAGCTCGAGGAATCAACTTAGTATTGAATAAAGGAATGATTCT
 GACACAATCTTGCACTTTTCATCATCTATCGTGCAACATATGAGTGGGCCT
 40 TTATTTATCATTTGTTCTGCTCTTTATCTTTTCTGTTTAGGATTATTCGTG
 CCGTCTCATCAGGATTAGCAGTACTATCTATGCCATCTTTGCGCCATTA
 GCTGATACAGTAGGTA

45 Sequence 3461

step.1003e05.cons.ok

CTATTTTCATGATGATCTGTCACTATAACATCTACACCTAATTCTTGTATGG
 TACTTATTTTCATGATGTCCCTGTATACCATTATCTACAGTTATTATTAAGG
 AAATCCCTTCGTCATAAGCATTTTAAATGCTAATTCATTAGGTCCATATC
 50 CTTCTGTAAATCTATTGGGAATATACCACCCACCTGAGCGCCAAGTAGAC
 GTAAAGTAGACACTAGAATTGTAGTAGATGTAACACCATCTGCATCATAGT
 CACCATATACTAGTATTTCGTTCAATTTGATCGATAGCTAATTTAATACGAT
 CAATGGCTTTTGCATATCACTTAATAACATATAATCATGACTAATGTTAG
 TGCCTTCAATAGATTTTGAATCTCCTGATTATCCACAATATTTTACTCT
 55 CTAATACCTTTTACTATAGGTGAAAGTTTAAATGATCACTGCTACTCT
 CACTTATAAAGGTATTTCGGTTGAGTTATATCCCAATTTATTTGATTAA
 TCATTATAATCGTCTCTCATTTCCACAGGTCAATTATACCGAAATTTTATTGA
 AATAAAAAGTACTTGCATTTTAAAAAGAATATATCATTTCAATCTTCAC
 AAAACTTTTATCATTTCAATTTGTTTCATTAACTGCTATCTTTATGATTAAAC

GCTAATATATAAAATCAAGAGGCAGGGGCATAATTCCTAGCAAAATAGACA
 GTAAATGAGTTTTTATAAATTCATTTACTGGCTTCTTTATTTACAATACCT
 CGTATTGTTGGCTCGCTTCTTAGGGGACAGCTTCAGCCTGTAGTCTTCAA
 CTTGTCCCTGCTCCCTCAATAGTCTCGCCAAAATACTTTGTGTTTTATATGTA
 5 ATTTACACATTGAAAATACTTTAAAAAATAAGGTCCTTCGTATAATTTAATAA
 ATAGCAATAAACTAAATTAACGAGGTGCCTTATGTATAAAGATTATAATAT
 GACTTAACTTACTCTACCAATGGAAATTTTCAGTTCTTATCCCCACACATGA
 TATTTTCACGACATGTAAATGATATTGTAGAAAATAATTCAGATACAGAATT
 CGATGAATTCAGACATCATCGTGGCTTAATATAAAAAATAATAGCTCAACG
 10 AGCTGAAAATAATCAAAAAATTATAAAAAAGACAATTTCTATATTATTTCA
 ATAGAAATTTCTCTTTAATTACTTATCTTGGACCTTTTGTTCAGCCTCTT
 GTAAATTTTTAAAAATTATTAATGTGAAGAAGTTAAACTAAGATTTTTTC
 ATCGTTAGATTCTTCTCTTTGTGTACTACTAATTTATTATTTTTTGTCTT
 TTTAAACTGTCGTTTCTTAAGCATGCCCATTAATGGTACAGCAATGAAAAT
 15 TGACGAGAATACACCAGATAATAATCCAATTAGTAATGCTAAAGAGAAAAT
 AAATATTGTTGGTGCACCTAATATTAATATTGCAACTACAACCTACAACCTAC
 AGTCAACACTGTATTAATAGAACGTGTCATAGTTTGTCTAATAGAGCGGTT
 GACTATATCATCAATTTGATCAGTATGCGTAATTACTTTAACTTTATGCAG
 ATTTTCTCGAACCGGTGAAAGTTACGATTGTATCATTGATTGAATAACC
 20 AACGATAGTTAATACTGCTGCAATAAATGTTAAATCTACTTCTAATCTAAA
 CAACTAAAGATTGCAATGATGATAAATACGTCATGTAATAATGCAAGTAC
 AGAAGATAGACCCATACGCCATTCAAATCTTAGTGAAACATAGATAATAAT
 GCCTATAGAAGCAAGTATTAAATGCAGTTACAGCATTTTGTAGCTAGCTCTTG
 25 GCTATGAGAGGTGAAACGGTATTAATTTGTGGATTATCTCCAAATTCAGA
 TTTTCACTTAGCACTTAATTTATTTGTCTTCTCAGCTGATAAATCTTTTTT
 AAAGTGAAGTGTGGCATTTTTATTTCCACTACCATTAATCTGTATTGTATC
 TGGTTCCAACTCAACTGATTTTACAGTTTTCTCAACCTGTGCTTGTGTTAT
 AGCATTTTTAGATTGAATATCTGCTCTTGTTCGGATGAGAAATCAATACC
 TAAGTTTAATTTAAATATTGAAATGATAATCAAACCAATAATTACAATTAA
 30 AATACTAAGTGAAATAAGTGGCTTAGCTAATTTAACAAGTTTAACTTTT
 ATATGATGTTTTTTAAATCATGTACATCTTTACCTTCATTAATATCATGTCT
 ATCCTTCTTCTTAACACCAATAAACCAGTATTGTTTTTTAAAGAAGTTTGA
 AGATACCAGTAATGATAACAACCTCTTGATAAGAATACTGCGGTTACAAA
 TATCATTAATAATACCTAAGAGTAACATGGTTGCGAAGCCTTTGACTGAACT
 35 TTCTCCAAAGAAGAAAAGCACAGCTGCAGCGATGACAGTTGTTAAGTTGGA
 ATCAAATATAGTTAAGAATGAACCTTTTATTTGCTTTTGAATACGCTTGT
 AAGCGTGCGTCCAATTCTTAGTTCATCTTTAATACGTTTCATACATTATGAT
 ATTGGCATCGACAGCCATACCTACACCTAAAATAATGCCGCCAATCCAGG
 TAGAGTTAATACACCTGATATGAAATTGAATGCGACTAAAGTTAAATAAAT
 40 ATAAGTGGTTAAGGCAATGATTGCAACTAAACCAGGCAAACGATAGAAACC
 AAGCATAAATAAATAAATTAATGCTATACCTACAATTGATGCAAAACATGGT
 CTTATCAAGAGCATCTTGACCAAATGTGCACCAACAGAGTTAGAGTAAAT
 TTCTTTTAAATCAACTGGTAATGAGCCGGCATTTAATAACTCAGCTATTTG
 TTTTCGCTTCTCAACACCTTTTTCCTTGAAGCCACCTGAAATTTCAAC
 45 ACTACTAGAATTAATAGGTTGGTCTACACTCGCTGCAGATATAAATTTAGG
 CTTTTTACCTTCTTGTGCTTTTGTAGCTTCTTCTTGTAACTATCGCCTTT
 TTCGAAATCTAACCAAACTACCATGACATTGTACGTTTTTTTAGAAATCTT
 TTCAGTTACTTTCTTAAATTTATCTTTACTTTTACTTTTAAATGTAAGTGT
 TGGTTGATTAGTTTCTTGTTTAAATCTTGTTTAGCAGAGCCTTGTTTAAT
 50 GTCTGAACCAGACATTAACATGATCTTCAGCATCTCTAATTGTTAAAT
 AGCTGTGTGCGATAATAATTACGCGCTTGTGCTTGATCCTTGATACCTGC
 TAATTGTACACGAATTCGATTTGGATCTTCGATTTGTATTTTCGGTTCTGA
 TACACCTAGAACATTTACACGATTTTCTAATGTTTGAGATGTAGCTTGAAG
 TGCTTTTTTATCAATTTTATCTCCTTTATTTAAAGGATCTACTTGAAGAG
 55 GACTTCAAACACCTTGCAAATCTAGACCTAAATTAACATTTTAACTAC
 ATTCTTATACGTCATACCTTTAGAGCACAGTGGCGATGATAT

Sequence 3462

step.1003e06.cons.ok
ATTTCTTAAAAATGTTGTAAGTCTTATGTTTGCCTTTTATTTTCACGAATCTTT
TTATTTTATTTCAAACAATTGTTTACATGTGATTTTAAATCATCTCTCTCTAA
ATACAATTAAGAATACTATACCTGTATCTGACATTTCGTCAATAGTTAAATG
5 AAACCTCTTTTCACATCGTCATGTATCATTATTATTTAGTAGACTGATAG
CATAATGTAGCAGTTTTATTTATTTAATTAAAAATATTAAATGGAAATAGA
ACGTAAGGTGACATTACAAGCAGTTAATTATAAATAATGTTGAAATCTCTC
GCCTCTCATCCGAGAAAAATGTATGTATATCGTCACAATGTATCTACTAT
TACACAGAACAAAATATTTAGTTTCAACAAGTTGTAGAAGCGGGACACTAA
10 TTATTTAAAAAGTGTTCATATAAGCACATCTAAATATGACGTAATGTGTG
CAAATTTAAAAATAACTTGAAATGTGATAATGACTTTGCTATATTAATAATTA
AGTTAATAAAGTATTGTTTCGTAGGACAAGTAACATAGATAGTTTCGATTTC
GAGAGCTTGTGTAAGTGAGAACAAAGTAATCGACATTCATGTGAATCTACC
TACTATATGTGAACAATCGGTAATAACCGTTATTTTAGTTAAGCGCAATTT
15 GAGGTAATAGTTTTAATTTACTTGAAATGTTAAATAGGGTGGCAACGCG
TAGACCACGTCCCTTGTCTGGGATGTGGTCTTTTTTTATTGTTTTATCACA
CGAAGTCATCCATAAAATTTGAATATATTTTATTTGGGAAAGGATGAAGGTT
ACATGTTAGACATTCGTTTATTTAGAAATGAACCTGAGAAAGTGAAGAGCA
AAATTGAATTAAGAGGCGACGATCCTAAAGTTGTCGACCAAGTTTTAGAAT
20 TAGATGAACAAACGCCGTGAATTAATCAGTAAACGTAAGAGATGAAGCGGA
AAAGAAATAAAGTGAGCGAAGAAATAGCTCAAAAGAAACGTAATAAAGAAG
ACGCTGATGATGTCATTGCTGAGATGCGTCATTTAGGTGATGAAATTAAG
ATATCGATAATCAACTTAATGAAGTAGATAATAAAATTAGAGATATCTTAA
TTCGTATTCTTAACCTTAATTAATGAAGATGTACCTCAAGGTGATTCTGATG
25 AAGAAAACGTTGAAGTTAAAAATGGGGTACGCCACGTGATTTTGAATTTG
AACCTAAAGCGCACTGGGATTTAGTTGAAGAATTAAAAATGGCTGACTTTG
AACGTGCTGCTAAAGTATCTGGTGTCTGTTTCGTATACTTAACATAAGATG
GCGCATTACTTTGAACGTGCTTTAATGAATTACATGTTGACAAAACATACAA
CGCAACATGGTTTACTGAAATGATGACACCTCAATTAGTGAATGCTGATA
30 CGATGTTTTGGAACAGGTCAATTACCTAAATTTGAAGAAGATTTATTTAAAG
TTGAAAAAGAAGGCTTATATACGATTCCAACGTCAGAAGTACCTTTAACAA
ACTTCTATAGAGATGAAATTATCAACCAGGTGTACTACCTGAATTATTTA
CAGCTCAAACGTCATGTTTCCGTAGTGAAGCAGGATCAGCTGGTAGAGATA
CTAGAGGGTTAATTCGTTTACATCAATTTGATAAAGTTGAAATGGTTCGTA
35 TTGTACAACCTGAAGATTCCTGGGATGCTTTAGAAGAAATGACACAAAATG
CTGAAGCTATTTCTGAAGAATTAGGTTTACCATACCGTCGTGTTATCTTAT
GTACTGGCGATATTGGTTTCAGTGCTAGTAAACATATGATTTAGAAGTTT
GGTTACCAAGTTACAATGATTATAAAGAAATCAGTTCTTGCTCTAACTGTA
CTGATTTCCAAGCAGTCGCGCAAATATCAGATTCAAACGTGATGCTGCTT
40 CTAAACCAGAAATTAGTACACACATTAAATGGTAGTGGTTTAGCAGTAGGTC
GTACATTTGCAGCCATCGTTGAAAACATCAAAACGAAGATGGTACATTAA
CAATTCCTGAAGCATTAGTACCATTATGCGGTGGCAAACTAAAAATGAAA
AACCAATCAAATAATCATATAAACTTAACCTAGATGATTACAATATTAAAAA
GTTATTTAACTGTGCGATAAGATTCTATAAACTTATCGACAGTTTTTTAATA
45 TATAAAGGCGAAATTGGTATTTTATATTCAATTAGGGATGGATGCTAAA
ATAGTTTTAAATGATTTTTCACGATTTAAATACCTGGTATAAAAGCGATTTT
ATAGATACGGTTGTCTGACGTCACACAAACGAATGTCTAAAAAATAGTGAA
TATTTATGTCTAGATTGAGTTATAAAATCATGAAAAATAGTACATGAAGT
AGCAAGTATTTATCTTTTAGATAACGCCTGTAAAAATATAAAATAGTGGA
50 ATTGATTTTTTAAACGTTATCATTTCAATTTCAAATATTAACTTAGTTTG
TCAAACAACGTTTATTTTTATATAACCATAAAATAATATCATAAAGAAGCT
ATATGAAATTCGAAGTATAATGCCAATGTTTCATATAGCTCCTATTCTAAA
TATTAATAATTAATAATTAATGTATCAATCGTCCTATTCAATTTGATGCAT
AGCCTGAGAGTGGGTAATGATCTGAGTAATCATTTGTATTTGTATGTTTTTC
55 CCCAAGATTTACGGACCATTCTGGTGACTTCACTCTATGTGTATCATTAT
GCCATGAGCTTGGTTGTGCATGGTCATGATCTAATAAAATATAATCTAAGT
GTTGAGGTTCTAATTTAGGATAATTATATTTTCGCAATACTATTACTGCTAG
TATCCCAACTTATGCAATTACCATCGAATTGAGTAGGTAATGAAACATTTA
AGTTATTTGCCATTTGTTGATATTCATCTGAATCTTTAATGACATTTAAGT

CACCACCGATATAGACGGGTTCATTTTTAGGGATATTCCTTGTCTTTGATAA
ACTGTTTAAATTTCACTCATTTGACTCTGTCTAATGTCTTTATCTTTTCCTT
TAAAGCATGTTGGATCTTCAGCTTGTAGATGTGTCCGATAAT

5

Sequence 3463

step.1003e08.cons.ok

ATGCGTTTATCAAACGTCTCATTTAAACAATTTGGTAAACCTCAAAAGGTAA
TTACAGATCAGGCACCTTCAACGAAGGTAGCAATGGCTAAAGTAATTAAAG
10 CTTTTAAACTTAAACCTGACTGTTATTGTACATCGAAATATCTGAATAACC
TCATTGAGCAAGATCACTGTCATATTTAAATAAGAAAGACAAGGTATCAAA
ATATCAATACAGCAAAGAATACTTTAAAAGGTATTGAATGTATTTACGCTC
TATATAAAAAGAACCGCAGGTCTCTTCAGATCTACGGATTTTCGCCATGCC
ACGAAATTAGCATCATGCTAGCAAGTTAAGCGAACACTGACATGATAAATT
15 AGTGGTTAGCTATATTTTTTACTTTGCAACAGAACCTATAAAAAACATGG
GCTTAAAGGGTTAATTTATATAGTATTACTGCAACATTGATTTGGGTCAA
AAATTAGGGGTATTAATAAAAAATGAATGAATCATAAATTTTATCAAGCTGATT
GGAGAGGTTAAATGCATTATATAAACTTTAGAGTATCATTAATATCTTC
ACCTCAGATATACAAAAACTACTAGCATATAAAGATATATCTTTTATAT
20 GCTAGTAGTTTAAAAGCTTAAACAATCAATCGTCCATTTTATTAATATATT
CTTTTTCTTTATAGTATAATTTTCTCGTCTTTCTCTCGTAGTTTTCAC
GTTTTTTTAAATCTTCTGGCGTATTAACTCTTTTGGCTTAATAAGATAG
TAGATTGTCTAGATGCATCGTCTTTCTGTACATATATGCACCATTACGGT
AGGCCGCTCTACTTATCAAAATGCATTCCAACCGGAGATGTAAGATTGATAA
25 AAAGTAGTGATAATAATAATCTGACACTGAAAAACCTGAATTCACAATAA
AATAGATCAGTACACCAACTACAGTTAGTAATACTGACAATGTAGAACTTT
TCGTTGAGCGTGACTTCTTAGAAAGACATCTTGAAATTTTACAATCCCTA
TTGCACTAATTAATGCAATAATACTTCTTAAGAAAAAAGTATCGAAGCAA
TAAGACTAACGATGTCTTTTATGATTTCCATTGAAGACACGTCCTTCCCG
30 ATGAATCTTGAAATTGAGACCGAACTGACAAACGAAATAATTGCAATTAAAC
ATGATTGAATCTAAATAAGACACTGAGTTAAAAATAACGCTCATCACACCA
ACAATAGACATAACAACAGCACTCGAGGCATCAAATGATACAACCTCTATCA
GCAGTAGTGGGACCTTTAATTAATCTGACTAGACAAACAAGTAGTGCCATA
CCAAAAATCACTAATGCACCTTATAATAAATATTTGAGTGAACATTTCAATC
35 ATCGTGTACCTCCAAAATTAAATCCTCATACTGCTTAATACTTTTATAGAA
GATTTCTTTATCTTTTCTGACACATCAATACTGTGAATAAAAAATTTAT
TAGTATTTTATAGAAATTCGAATAACTGTGCGATCCTGGAGTAATAATAATTA
AAATCGTTAAAAAAGTAATAGCCCAATTACTTTTAAATGAAGTTTCATATG
TGAGTAAACCTGGATTAACTTCATTCTGCTTAAATAAGATGTAATTTATGG
40 TACTTATACTAGAAGTAATAAGCTGGTATAGGTATACAGCTAAAAATTTAA
TAGCCACCCATATCTTTTCAAATAAAATTTCTTACCAAAAAAGCGATGCA
GAATATAAATCACAATTAAACCAATTAAAAATCCAGCAAAGAAGGTTGTAA
ATTTAAATTCATCTTCATCTTGAAAGAGTACCCAAAGGAACGCGATAACAA
TATTTAATACAACTTGCTTCACCTTATGATTAACCTCCTTCAAATGTGTATT
45 TACATCTTTTGAAGACATCTTCTTTCATATTAAGATTTGTTGCATCCTC
TGTTACTTTCAGAACAACAGGGGCTGCAATACCCATCGCTAATACCACTAC
AACTAAACACTAAGTAAACCTTTACGATAAATAGGTAGTGGTCTAAATTG
TACTTGTCTCCGTCAGCATCACCAAAATACATTATAAACATCACTCT

50

Sequence 3464

step.1003e09.cons.ok

GATATTGAAAAATAGGCAATAAAGTTATTAAAAACATAATAACTAACGTAG
CACTATGTTTAGACATTATTTTAAACCAACACCCCTCCTTTTCATATAAA
55 AATTTTGCTCCGTTTTTTCATATTAGACATTATCACTCATCTTAACAATAAC
TATTTCTAATTTTTTCGAATTAGACACAAATATAAAAAATAGAAATTTTATCT
AGTAAATTAATACAAAGTATATATAAATCATAAAGTTAATAAATCAATTAA
GTATGTAGCAAAATGGCAATTGTGAATATAAAATAAAGTTTATTCATTTTA
GAATCACCTTACTAAGTGATCAACAAAAATTCGAAGATTGAATAAGTTGTA

TAAAGTTTATTATAATTAACCATACTATTAAATCATAAGTAATCTGCATAT
 GCCTTAAAGATAAAATTATATATTTAATGATTAAATGTACAATCTAAAGTTAT
 TAGTAGCTGGATTGATAATATAGCAAAATTCATCTACTTTTCTTATACCT
 AACTATACAATAGGAATGTGATTCAACTTTATATATCATATTTAGGCCAAG
 5 AGAGTATGATATTTAAAAATTATGTCCGAAGTTTAAAGAAGCTTGATGCCAC
 AAATTTAGAAGTTTAAAAATTATTTAATTAACTTTAGTATCATCCTTATTC
 TAGTGGACGCTGAAAGGAACATTAATTATGATTACAGCTTATAAACACGC
 ACTTCATCATCCTTCACAAATTTATGAAACTGAAATCAATCATAGTGCTTC
 ATGGATAAATGTAGTAGAACCAGATAGAGAAGAAATGAAGGTTTGATGGA
 10 ATTTTACAATATACCTGAGGATTTTATTTCGAGATCCGTTAGACTCTGAAGA
 GAGTGCACCTATTGAATATGACGAAGATACGGGATACTCGCTAATTATTAT
 TGACTTACCTATCGTCAATTCACCAATCGGCGTGTCTTGTCGTTTGTGAC
 CATCCCATTAGGCATTATTATTGGAATGGTATCGTTATGACTGTTTGCGA
 TGCTGAAAAAGAATTTTGTAGAAACTTTGCTAAACAAGAAGATATTAACCT
 15 GAAATTTACAGTCGATTTGCACTTGAAATACATAACAACATAGCAATCA
 CTACAATAGAAATTTAAGATTGCTTAATAAATCTAGGATTCGTATTGAAAG
 AGAACTCAAAAATAACATTACTAACAAGCAGCTTTTAAACTAATGGAAAGT
 AGAAAAAAGTTTAGTATACTTCTTAGCTGCACATAAAGGTAACGACACAAT
 TATTAaaaaactcTTTCGTCTTCCTGCAATCAAACGTTTGAAGAGGATGA
 20 AGAGCTACTTGAAGACTTGGTGATAGAAAACAACCAAGCTATTGAAACGAC
 TGAATTACATCAACGAATACTTGAGAGTATCACTTCTTCGTATGCTTCATT
 ATTATCTAATGATATGAATAACATCATGAAAACATTAACGTTGTTTACGGT
 TCTTTTGACCCTTCCTACACTCGTCTTTAGTTTCTTTGGTATGAATGTACC
 CTTACCAATTGATGATCATAGTTACGTGTCTTGGATTATTGTTGTGGGAAT
 25 TTTCACTTATTCTAGTAGCTATCGTTAGTATCTTTTATGAAAAAACAAAA
 ATTATAACAACAATAACCCTTAGTTATAGATTAGCACTCAATCTAACTAAG
 GGTATTATTATATGTTTAAAAATTGTGTCAAAATTTAAGTCTCTACTATGA
 TTCAACAGCTGTCTTATTTGTTTCTTCGAGAAAGTACTAAAGTGACGAC
 AACTGAAACAACAATGCAATTAGCATTCTTATTAAGTAATGTAACCAACC
 30 AGCGTGTGTAGGATTTATAGATATAAATCCAGGTAATCCAGCTGTACCAAG
 CGCTATCGCTTTTACTTTGAAGAATGAAATATAAGCCGCACCTATACCTGA
 TCCTGCAACAGCACCTATAAATGGATATCTCAATTTAAGATTGACACCAAA
 CATTGCTGGTTCGTGAATTCCTAGTAAAGCTGAAATACCCGCCGCGGAAGC
 AACACCTTTTAATTTTTTATTTTGCTTAATGATAAAGAATGCAGCTAAAGC
 35 TGCACCACCTTGTGCAATATTGACATCGTTGCGATTGGGAAGATAAATGA
 ACCACCTGTTTTAGTCGCATCAGCAATTAATGTCGTTTCAACTGCAATAAA
 GCTATGGTCGATTCTGTAAATGACGATTGGCGCATATAATAAACCAAAATAT
 AAGTCCACCAATAGCTCCACCAAAATTCATATAACCAAGTCAATCCATCAGA
 TAACCAATAACCTAATTGACGAGTGACAGGCCCTACAAATAAAAAATGTTAT
 40 AAATGCTGTAATAAAAAATTGATAATAATGGCGTCAATAAATTATCTAACAC
 AGTTGGAAATACTTTACGTAACCATTTTTCAATCGTAGCTAAATATATGT
 TGCTACGAGCATAGGTAATACTTGTCCCTGATAACCTACTTCATTAATATG
 TAGCCCAAAGACATCCAGTGTGGAATAGCTTTTCTTCTTCTAAAGCTTT
 TGGAAAATCATATGCACTCATCAATCCAGGATGAACAAGTATCATACCTAG
 45 AGCTGCACCTAAATAAGGATTACCACCAAAATCGCTTAGCTGCACATAAAACC
 AATAAGTATTGGTAATAATGTAAAAGGTGCATTTGCAAATATATTAAATCAT
 ATCAGCCAAGCCAGAAAATTGACTATGTACATCTATGATAGATTTACCATC
 ATAAAACAAATCTTTAGCTGTGAAAATATTATTTAATCCCATTAACAAACC
 ACCAGCTACAATAGCCGGAATAATAGGAACAAAGATATCAGATAACATTTT
 50 CACAAATTTTGGAAATGGATTATATGCTTAGATGATTTATCTTTAACTTC
 TGAAGT

Sequence 3465

55 step.1003e10.cons.ok
 CTGATGCAAAGCTCATTGAACACTCACCCCTTCCTAACCCCTTTTTTCTATTCT
 TTATTTATAGTATCGCCAATGAGGTGTAAAGTTATTTTCTAATTCTAAACA
 CACACTTAAGCTATCATATCATAACTACGAATTCGCTTATAATTATTTCTT
 TTTATCACTAGGAGTAAAGCGAATTGTACTTGTTAATTCTAAGGCAAGTTC

ATAAATCATTTTTGATAATACTTTTGTGTTATGTCTGACATAGTGTTCATT
 AGATATTTCAACTAAATTAGATGCCGTTAAACCTCTAATTCACATCTTTT
 TAATTGTTCTTTATGTACTGCTACTGGTTTCGAATTCTTTCTTCATATCG
 TTGTAAAACATCTTTACTATAGGATTCTGAGCTACATATGACAAAATCAAT
 5 AAATGGTTACCAACTTGTGAGTAAGTGCATCAATATGCTCTTTGACATC
 ATAATTATCAGTCTCGCCTGGTTGTGTCATAACATTAGATACATAAAGTTT
 TGGAGCAGATGTACGTAATAATGCTTCTGAAATACCTTTGACACATAAATT
 TGATATAACACTTGTATATAATGAACCTGGTCCTAAGACAATTAAATCTGC
 TTGTTCTAAAGCTTCTATCGCTTCATTTCATTGGTTCAACATCACTTGGTTC
 10 TAAAAACACACGATCTATTTTTTTATGTGTTTTAGGTATATTAGTTCTCC
 ATGTACAATTTACCCGCTTCCATCACCCGCTTGAGTTGCACACTTGCGTT
 TGTTGAAGGGATGACTTGGCCTTTAATATTTAAACCTTTGCTTAACTCTTT
 AATAGCGTGTCCAAAATCATTTAGTAATGTTAGTCATTCCAGCAATAACTAA
 ATTACCTAATGAATGCCCATCTACTTGATTTTACCACAAACGGTACTGGAA
 15 CAATTGAGTTAATATCGATTCTGAGTCACTTAAAGCAGCAATGACATTACG
 AATATCACCGCGCTGGAATATCCATGACATCTCTAATTTTCCCGTGCT
 CCCACCATTTGTCGCTACAGTAACAATGGCAGTAATGTCTATTGGAAATTC
 TCTAAGGCCCTCTAGCAAGGACAGAAAGTCCAGTGCCACCACCAATAAGTAC
 TACATTCATTGTGTTTCTATTCTCTCTCCACTTTCAATATGCGCATCTCTA
 20 TGATGCACATAAACATTATATTTCAAAAATCTCGTTAAGATATTCAGCTAAA
 CGTTTAGCTAATGCGACTGATCGATGTTGTCCACCCGTACAACCTATAGCA
 ATAACCAATTGCGATTTACCTTCTTTTTGTAGCCAGGAATCATAAATTTT
 AATAAATCTGTTAATTTATCAAAAAATATTTGGGTTCTTTCCACTTCATA
 ACGTAATTGTACACTGGCTCATCTAAACAGTAAATGGACGCAATTCCTCT
 25 ACATAGTAGGGATTAGGTAGAAATCTGACATCAAAAATCAATCAGCATCC
 ATTTGTATACCATGCTTGAACCGAACTTGTACGTTGATTGTAAATGTT
 TCAAAGTTTTCATCTAAATAAACTTTGAAATGCGTTGCTTTAATTCCTTTA
 GGTTTTAATTTTGTGTTATCAATCACGTAATTAGCGATACTTCGGATTTCT
 GATAGATGTTACGTTCTCTCATTTTATGTCATCTATTAATGATCTTTGTCTT
 30 TGTTTCAATTTAGTGGGTGCGCTCTTCTTGATTCTTTATAACGTGAAATAATT
 TTTTCAGTTTTAGCTTCTAAAAACATAACATCTAAATCACGTCAATTACGA
 CTTTTAATAATATCAATTTCTTTAACTAGAGATTTAAATAATTCCTTACCT
 CTTAAATCTATTGCAATTGCTACTTTTGGCAATGAAGGTCAATGTGTATTA
 CTGGTATGCAATTTGGCAATGAGGTAGCTATGCATGTGAGCAGAGGTGACG
 35 TCTTCGGTATGATAGGAGTTTATGTAGTATACTAGCAACCTGGGGATTAC
 TACAAGCTACACATATGTGGCTATTGAGTATTATTGGTGGTTTTGTAGCTT
 GGTTCATTAGTGGGTGATTATTTTGGAAATGTGGAATTCATAGCACATA
 AAAGAAGGGATAAACATGGCTGGAAGACCAAAAGATCCAACAATCAATAAA
 AAAATTTTACTGAGATTCAACGTTTATTGGAACGACTCATTTTAGAGAT
 40 ATTACTATAGATCAAAATTTCTGAGAATACTGGTATTCTAAAGCAACTATT
 TATAGACGTTGGAAGGATAAATCTTCAATTATCATGTCCGCGTTTATTGAA
 CAATCTCAATATATTGCGATTCAATCAAGATAATTTATATGATGATTTA
 TTCCAGTTTTTAGTAAAAATAAAAGATATCTATAAAACAAAACCTAGGTAGT
 GCTGTGATTGAAATATTAATTAGTCATCAACAAATGGAAGCTAGAGAACT
 45 TTTATGACTAATTACTTTAATCATAATCGCAAAGTTTTAAAGAGATTGTT
 CGTAAGCACATACAAGAGGAAGAACAAGATTTGTTTATTGATTTAATCTTC
 TCACCCATCTATTTAATATATTAATTAAACCTGAACTCTGGATGAAAAT
 TACATTAAAAAGATGTTGAATCAAGTTTACGTATATATCATTGATATGAA
 AAACGCTCCTTTCGAGGAGCGTTTTTTGTTTTTCATGTAACAGGTTAGCCT
 50 AAGGTCAGTCTATTAACCTTTAACCATACTATGTCAAAGCATTTTGACA
 ATGATAAAGGTGACATCGCAAAAACCTATTGATAGCTTTTATAGAGAAAAC
 CTGATTGATTTAATAATAATATTAATGATTTCAAATCATAATTAAGGAG
 TTTAACGTAATGAATGTTAGCAATCAAATCAAAAAATTTAGAGAAAGAGAT
 GGGTATTACAGAATTTTGGCTGAGAAAATGTTGTTTCAAGACAAACC
 55 ATTTCAAATTCGGGAAAATGATAAAGTTATCCGGATATTCATAATTTACTT
 ATTATGTGTCAACTCTTCAAGGTATCTCTAGACGAACTAGTGGAGGATGAT
 TTGAAAAACGTGCAAAATCAAAAATATTAAGAAAGAAATAGACTTTTGGACA
 TGGATGATGATAATTCCAATGGTATTAGGAAGTGTCTCATAGGTCCATG
 ATTCATTACTTTTCATTGGTTAGGAATAGGAATAACATATTTAGTATTTTAA

ATAGGTATCATTTGCTTCTACAAAATTAGAAAAAATTAAAAATCAAAAATGAT
ATTGAAACATACGACAGGATATCTTTAGAGCACAGTGGCGATGATATC

5 Sequence 3466

step.1003e11.cons.ok

TATAAATTTAAATGTACATTAGCAAATGCATCGTTATCAAGTCCGTCTACA
GGGTCACCTACTACGACAAGTCCGTGTCGGTATTATCTGATATTAAATCT
GATAAAGTAAATTTGGAAACCAATTTTATATCTTGCACTCGGAATTTCA
10 ATGCCCTGGCGCATACGGGTATGATATCTAATTGTTTCTAAATCAGCATCA
TAAGGTAAGTCTTCCTGCACATATAAGATAAATTTCTGGTCTAGTAATGGT
GAAAATAATTTCTGAAAGAGAGACTGAGGCACCATCATTAACAATTTGGTTA
GATAAGAGTGTTCATATGCAGGTTCGTAGTGTAGCAATTGCTTGCGTT
GCCCTTGCTAGTCATACTAACTTTTATACCCTGTAACAGTCGAACGGTCTTTG
15 AAAGTTAATTTGGTCAATTAGTTCCTTCTGTACATGATATGCTTCTTCTTCG
TTTAATTGATAGTTCCTCACTAATGAAATCGATGGGTTTTTATATCTATAA
GCTTTAAATAAAACTTTTAGCAACTTCTTTATTTGTTAGTGTATATTGTAC
CCCTCCCCATAAATTATCTGAAAATTTATACTATTATGATAACTTTTAAATG
TGAAGTGTGCAAAATTGATTTAAATATACTTCATTTAGGGTATAAAATAATT
20 AATAATGTAAATTTTGATAGGAGTGGGTATAGTGTGATGATTATGTTTATGA
GTTGATGAAACAACATCATTCAGTTAGAAAAATTAAGAATCAACCACCTGG
TTCTGAAACGGTAGAAAAATTAGTAGAGGCGGGACAGAGTGCTTCTACATC
CAGTTATCTTCAAACCTTATTCATTATTTGGTGTGGAAGATCCAAGCATTA
AGCGCGTTTAAAGGAAGTGTGAGGTGAGCCTTATGTTTAGATAATGGTTA
25 TTTATTTGTATTTGTTTGTAGATTATTATCGTCATCATTTAGTAGATGAAGT
TGCGGCGTCAAAATATGGAGACATCATATGGTTCTGCAGAAGGACTATTAGT
AGGTACAATAGATGTTGCATTAGTTGCGCAAAACATGCGAGTTGCTGCCGA
AGATATGGGGTATGGAATTGTTTATTTAGGGTCATTGCGTAATGATGTTGC
GCGAGTGGGTGAAATTTTAAATTTACCTGATTATACGTTTCCGTTATTTGG
30 TATGGCAGTAGGTGAACCTTCTGATGAAGAAAATGGGTACCTAAACCGCG
CTTGCCATTTAAACATATTTTTCATAAAAGACCAGTATGATGCGAATCAGCA
TCAACAACGTAAAGAATTGGAAGCATACGACCAAGTAGTGAGTGAATATTA
TAAAGAACGTACTCACGGTGTGCGTACAGAAAATTTGGTCACAACAAATAGA
AACATTTCTAGGACGTAAACACGTTTAGATATGTTAGATGAATTGAAAAA
35 AGCAGGATTTATTCAAAGATAAAAAGAACCCTGAGACAGAAATACATGTC
AGGCTAGGGTGGGGCGATATTTTCAACAGCAATGTTGCCCCGCTCTTTT
TAATTGTATGTCGTAATGATCAATGACCACTTTTCTTACATTTAATTTGCA
GATAAATCGCCGTAATCATTTGAGTCAAATGTTTTTTTGTCCCAATGATTC
GTTAAGCGTGCAGTACCTGTTCCAGCTAGCATTGAGTCATTAACGTTAAGT
40 GCTGTACGACCCATATCGATGAGAGGTTCGATAGATATCAGTACCCCTGCG
AGAGCAACTGGTAGATTAAAGTGTAGATAATACGAGTATTTGATGCGAATGTT
GCCCCGCCACCTACACCTGCAACGCCAAATGAACTTATAATAACAACAGCA
ATAAGTGTAAACAACAAATTTGGAAGTCAATTTCTACATTTGCTACTGGTGCC
ACCATAACTGCTAGCATAGCAGGATAGATTCTCTGCACAGCCATTTTGCCCT
45 ATGGATAAACCAAAAGTTGCAGAGAAGTTTGCAATTCCTCAGGTACACCT
AAACGTTTTGTTTGCCTTTGAACATTTAACGGTAATGCACCTGCACCTTGAA
CGTGAAGTAAATGCAAAGATTAGTACTTCTATTGTCTTTTTACGATTTTA
ACGGGATTGATACCTAAGACACTCAGTATAATTAATGGATAATATACATT
GTGATTAGAGCTGCGTATGAAGCAATTAAGAATTTACCTAACGTCCAAAT
50 GCAGAAAAATCACTTGTGCGCAAGAGTAGAAGCCATAAATGCAATGCA
TAAGGCGTTAATCGTAAACAAAAGTTACGATAGCCATAACGATAGAATAG
ATTGCTTCTATACCAGTTTTAAGTAAGCTTCCATGTTCCGGCTGTTTTCTT
GCAACTCTAAGATAAGCAAAGCCACAAACGTTGCAAAAATAACAACGCA
ATTGTGCAAGTTGTACGTTGCTCTGTGAAATCTAAAAATGGATTGCTTGGG
55 AATACCTCGAGAATTTGTTGTGGTAAAGTGTTCGAGTTAAATCTTTGGCT
TGTTTTGAAATTTCTGTACCAGTGAATGTTCTGCACTACCTAAATCAATA
GACGATGCATCTAAACCAAGATCAAAGCGTAAAAAATTCACGATAGCT
GCAATGGCTACAGTACCAATTAATAAATAAATAAAGAACCGATCTTA
GCGAATTTTTCACCAATTTGTATTTTGCTAAAAGCGGCAACAATTGAAATG

5 AATATTAGTGGCATGACAATCATTTGTAATAGTGCAACATAACCATCTCCA
ACAATACTAAACCAGTCTGTTGATTGTTTCGAGAGTTTACTCTGCACCA
TATATGAGGTGTAATACAATGCCAAATACGATACCTATTGCTAAAGCGGTA
AAGACACGTTTAGGGAATGAGACGTGTTTTCTTGCCATAATTTAACATT
ATTAGGAAAACAAGTAACACGATGACATTAATTACTGTAAAAAGTATTTCC
ATATCGATA

Sequence 3467

10 step.1003f01.cons.ok
TATAAGATAATATAAACTAAGCCTGAGACACCGATGAGTAATCTCAGGCTT
TTTTACGTATTGAAAGTGATTGAGTAATTAAATAGGGTGGATATCAAAAAT
GATGTTGGTCGATATGTTAGGTTCTTTAAGAGAATGCAATGCCAGGCTCAA
TATCGACGTAGTTATGAAGCGGTGTGTATTTAAATGGCCGTTAGTATCAT
15 GTATAAACGGACTATTGGAACCTCAACTTTTAAATTTGATATAGTATAATAT
AAAGGAAATTTAGAGGACGTGACGTTTCGATATTTGAATAACTGTATTAAG
GAATATCGAACTTTTAAATATCTACGCAATTTAGGATAAAAAATCAACAT
TGAGTGAAAGCCAATTTAGTCGAAGCATCTCAATAAAGCGTAAGACACTCG
AGTTCCCCCACCCCCATAAAACAAAATAACTATGGTACATATGAGAAAGG
20 ATGAGTCGCTTGATACGTATGAGTGTATTAGCAAGTGGTAGTACAGGTAAC
GCCACTTACGTAGAAAGTGAGAAAGGTAGCCTACTCGTCGATGTAGGTTTA
ACTGGCAAGAAGATGGAAGAACTTTTCAGCCAAATCGACAGAAACATTAAG
GATTTAAACGGAATTTTAGTGACACATGAACACATCGACCATATTTAAAGGT
CTTGGTGTTTTAGCACGTAAATATAAACTTCCGATTTACGCGAATGAGAAT
25 ACATGGAAAGCGATAGAGAAGAAAGATAGCCGATTCCTAATGGATCAGAAA
TTTATCTTTAATCCATATGAAACGAAATCTCTTGCAGGATTTGATATAGAA
TCATTTAACGTGTACATGACGCGATTGATCCACAATTTCTACATCTTCCAC
AATAACTATAAGAAATTTACGATGATAACTGACACTGGTTACGTTTCAGAT
CGTATGAAAGGTATGATTCAAGGTAGTGATGTCTTTATGTTTGAAGTAAT
30 CAGATGTCGATATGTTACGCAATGTGTGCTATCCATGGAAGACGAAACAA
CGTATTTTAAAGTGATATGGGTCACGTATCCAATGAAGACGCGGGTCTTGCG
ATGAGTGATGTCATTACAGGTAATACGAAACGTATATACCTCTCTCATTTG
TCACAAGACAATAATATGAAAGACCTCGCACGCAATGAGTGTGGACAAGTG
CTCAACGAACACGATATCGATACAGAGAAAGAAGTATGCTTTGCGATACC
35 GATAAAGCACAAGCCACACCGATTTATACACTATAATTGAATAGAGGTAAG
ACAAACGACCTACTTAAGTATTACGCTTAAGTAGGTTATTTTTTTGAAAAA
GTGATCAGGAGAATTGACAGGGAATGATTATGTGTTAGTTTAAACAGGCTTG
CGCATGAACGATTGATTAAGGAATTTTTTTGAAATGAAAGCGCATCATGGA
AGGGGGAGAACGATGTCTTACACGCTTTGTGATTTATATATAATAACATTC
40 TTTTACTTTTAAACTGCTTAAATACCGTTCTTCTGCGACATTTGCACGCT
GCGCAATTTCTTTAGTACTCGTTTATCAAATCCTTGTTCACTGAAATAATG
CAATAGCACTTTCTATGACACGTCTTTGTTTATCGCTTAAATATTCAAGTT
GAGGCACAATGGTTTCAACTAATGACTTAATATCTTGGTTTATATTGACCT
CCTAAACTTTTCTATATCTTTTCATGCCTATAATTTAAATCAATAAGA
45 AAATGAAAAATGCGAATAACACAATTAATAAATGTAAATATCGTTGAATC
CATAACCCTTGATCATAATATTTTGCATCGTTTGGCCGGTATAGAATAACG
GCATGATATGTGAAAAGTATTGTAATCCTTTATTCATTGATTCAATTGGTA
TAATGCCGTGCAAAATAGTACTTGTGGCACTATGACTAATGGTATAAAATTGAA
TCATTTGGAAATTCGAGGAAGCAAAGGTAGATAATAATATACCGAATGTCA
50 CAGCGACAAGCGCTGTTAAATATGCGCTTAATAGTACGAACCATATCGAAC
CTACTAAGTCTATATGCAGAAATATAAATGTCATATAATACGACAACATTTG
TTTGGATAACGCTAAAACCTACCATAACCGAAAACATAACCAAAAAATAATTT
CACTTCTTTTATTGAGAGGCAAGTAAACGTTCTAATGTGCCAGAAGTAC
GCTCTTTTAAATAGCCAATGCCAGAAATTAACAAACGTAAGAAAAAGACAA
55 AAAATCCAATTAATAAGGATTTATCATATCAAAATACGTAGAATCTGATG
AACCATATAAATAGTGCCTTGTAGTTTATATGGTTTAGCCATATCTTGAT
GAGGCGTGCTCCCCCATCCCCGGGCATTTTTTGTGTATTTTAGTTAAACG
CTTGATGCAATTTATTAGTATTATCTTTTCATGGCATTCATGTTATGACTCA
TTAACCATTTTGTATTTGCACCTGTTAGTTCTCCTGCTGTGTAGGATTAT

CGTTAGCATAAGTCACTGATACTTTTTGACCATCACTGTGCAAAAAATCCTG
 TTAATTTGTCGCTCTTAATTTTATCACTTATATCATGTGCATTTTATAAT
 GTTTAACGTGAATATCTTTATCATGTAATTCAGTCATTAATGAATCTGGTA
 CATCGTGAACCCCAACTGTTACGCCATTCGTATTATCGGCAACTGTAAAAA
 5 TGTAATATAGTAGCGACAATAGTAATATAGGTGCAATAAGTAACAGTGCCA
 ATGTTCTTACATCTCGAATCGTCTGTTGAAAATACGTCTGCTATGTGCA
 ACGAATTCATTTATCCCTCCATATTCAGAAATACGCCTTCAATCGTATCT
 GTGTGATATTGTTTTTTCAC'TTCATCTGGTGTACCCGTGCGCAATATCTTT
 CCTTGATTTCATTAATACGAGCTTATCACAGCGTGTAGCTTCATCTAATACA
 10 TCGCTTGTCACTATAATGCATTTATCTTCAGCCCTTGCTTTAGTTAAATCC
 TTCCAAATCTTTAGAGCACAGTGCGCATGATAT

Sequence 3468

15 step.1003f02.cons.ok
 TACAGAAAATAAGACCTCTAAATGCTAACTAGTTAGTATTTTGAAGGCT
 TTTTATAAAAAGAGAAACATCGAAGGGAAGATTCATTCGACATTTCTCC
 ATAATGATATTAATTAAGCTTTTATGCTTAGTTGTATTGTTAATTAATTAC
 CTTCTCCGTAGATGTCAATCCATTCCTCTCTTTTATCTAAAAATGCTTGAG
 20 CAATTGTTTTAGCACCTTCTAAAGAAATGACTTGCAGCCCAACCACATGAA
 CCTCATTACAGCTGGGACTTCGCTAGCATTTAACACATCATGCAATGTTT
 GATCGATAATATTTAATACGTCAATCGTAGTCGTATGATTAATAAATGAAA
 CATAGAATCCAGTTTGACAACCCATAGGACTTAAATCTACTACTTTATCAG
 TATGATTTCTAATGTTTTCTGCCATTAAATGCTCTAAGGAATGTAGACCAG
 25 GCATATCCATATGTTTCCTTATTGGGCTGTTTGAACGAATGTCATATTTGT
 GTATGACATCACCATTAAAGACCTTCCATAGTCCCGGCTAGACGAATAAAG
 GTGCAACAACCTTAGTATGGTCTAAATTAAAGCTTTCTACATTCATTTAG
 TCATGCATTTATCCTCCTTATTGCTCTACAGCAATTTTGTTTAATTCAA
 TAAATAGATTGTAACAATACCTCTAATACATTTACAATTTTCTAAACTAAT
 30 AGACTGTTGTCTTTTAAATTTGTAAGAAAATTGTGTAAAAATTAATGACAAC
 CAATATTCATCCGATAAAATAAGTAATAATTAAGGAATGATTGATAC
 GCAAAGTTTCAAAGGGAGTTAGATACAATGCATAATAAACAAAAGATATTA
 GATTTTATAGAAAATAATAAATATGATTATGTTGAAATAAGTCATCGTATT
 CATGAACGCCCTGAATTAGGCAATGAAGAAATTTTGCATCGAGAACATTA
 35 ATTGACCAATTAAGAGCAAATCGATTGCAATCGAAACGGATATTGCAGGA
 CATGCAACAGGATTTATAGCAACGTATGATTCTGATATGACTGGACCGGTT
 ATAGGATTTCTAGCTGAATATGATGCTTTACCTGGTCTTGGTCACGCATGC
 GGGCATAATATTATTGGTACTGCTAGCGTACTTGCTGCAGTAGCACTAAAA
 GAAGTCGTCGATGAAATTGGTGGTAAAGTAGTCGTTTGGGATGTCTGCT
 40 GAAGAAGGTGGGGAAAATGGCTCCGCAAAAGCTTCTATGTTAAAGCAGGT
 GTCATTGTGAAATTGATGTAGCATTGATGATTCATCCTGGAAATGAAACT
 TATCGTACAATTAATACTTTAGCTGTGGATGTTCTAGATATTAATTTCTAT
 GGACGTAGTGCATATGCATCTGAAAAATGCAGATGAAGCATTAAACGCTTTA
 GATGCAATGATTTTCATATATTAATGGTATAGCACAGTTAAGGCAACACATT
 45 AAAAAAGGACAACGAGTTCACGGGGTATTTTAGACGGTGGTAAAGCGGCT
 AATATTATACCTGATTTTACACATGCGAGATTTTACACTCGAGCTACTTCG
 CGGAGAGAACTTGATGTTTTAACTGAAAAAGTAAACCAAATTGCAAGAGGT
 GCTGCTATTCAAACCTGGGTGTGATTTTGAATTTGGTCTATCCAGAATGGT
 GTAAACGAATTTATCAAAGCACCTAAACTTGATGATTTATTGAAAAATAT
 50 CCAACTGAATTAGGAGAAGAAGTGATTGATGATGATTTTGGCTATGGATCT
 ACAGTACAGGTAATGTAAGTCATGTTGTACCAACAATAACATCCACATATT
 AAAATTGGTTCTCGAAATCTTGTAGGACATACCCACCGCTTTAGAGAAGC
 GGCTGCAAGTTTACAAGGTGATCAGGCACTCATTGAGGTGCAAAAAATTTT
 AGCATTAATGGGACTAGAATTAATCGAAAAATAAACCGTTGTTAGACGAAAT
 55 AATTGAAGGCATACGCATATAAAGGGGCATGTTAAGTAATGAATAATAAA
 GTAAATGGTCCGTTTCTCACACTTAGTGATTTGTATAATGATGACATTGTT
 TACACATCTCGACCTTCATATGTGTCGAACCCATGGTTGAAACCAGATGAA
 CACCAATCAAATTTCTTAAC'TGGAAGAGAATTACTTATAGCTAATCGTTTT
 CCTGTCAATTGTTTCATGAGGCAAGTGTACAAACAAATTAGAACAACCTTTTT

AATATAGTGGGTAGGGAAATTCACCACATACCTTATAAAATTTAAAGATCAA
 GAAACATATGAAAGCTTGATAAGGAATTTAGCTCTTCATCAAGGTAAAAAA
 ATCTACTTTCAATATATTCATGATGAAGATATTTACCTAAAGAATATTAT
 GCACCTTGATAAAGATGTTTTGTTGCTCTTAATAATAAAGCACGAATTTCCA
 5 GAATGGACTAATAACAAATATCTACCACAAAGAGAAATGTCTCAATTAAA
 GATTTTGAATCTCATATTCAGCATGGTCGTATCCATTTGTCATAAAACCA
 GGCGATGATTTACCTACAGCAGGAGGATATGGTGTATGATTTGTTATAAT
 GATACAGATTTAGCTAAAGCTATCACACGCATCAACAATGCATCAGCAGAG
 ACTGAAAAATTAATCATTTGAACAAAAAATTAATGCAGTGAATAACTATTGT
 10 GTACAATTTGCTTATTCAGATGATATTGGTATCAAATACTTAGGAACAGCG
 CAACAGTTAAACTAAGTACTATGGATTTTACAACGGAAATGAAAATGTTAAT
 GATGTGCCCTCAGAATGTAATAGACGCTGGTAGAGAGATTATGGAAATAGGC
 GTAAGCAAAGGTTTTTTTTGGTGTAGCAGGTTTTGACTTACTAGTAGATGAT
 AATAATGATGTTTATGCGATTGATTTAACTTTAGGCAAAACGGATCAACG
 15 AGTATGCTACTTTTAGCAAAAAGATTTAACTCATGGATATCATAAATTTTAC
 AGTTACTTTTCTAATGGAGATAATACAAAATCTATAATGCTATTTTAAAA
 TACGTAGAATTAGGTGTACTTTATCCACTTTCTTATTACGATGGAGATTGG
 TATGGAAAGAATCAAGTTAATTTCTAGATTTGGCTGCATTTGGCATGGGGAA
 AATAAAGAATTAATTAATCGATATGAACAACAATTTATATTGGAAGCTGGA
 20 TTATA

Sequence 3469

step.1003f03.cons.ok

25 ATGAAGCAGCCGCTGGATCAAGATTTTCGGGAACTTTAACTGCATAATTTCG
 CATCAACAATGACTTCTTCAGCCATTGCCCATCAACAGTATATCCGGCAT
 TAAGGACATCCCTGCATAATGTTTCTCTACCAGTGTGCAATACTCACAAT
 TACCACATGCTGCATACATCCATGCTATAGATACACGATCCCTTACCTTTA
 ATGAATCTACGTTATCAGCAATTTAATGACACGTCCAATTCCTTCATGTC
 30 CTAATGTTACACCTGTTACGTCACCGAAATCTGCATTTTAAACATGTAAAT
 CAGTATGACATACACCACAATATTCAGTCTTTACTAATGCTTCTCCAGGTC
 TTAATTCCTCTTAAACGTTTTTCTTCAATACTTACTCTATGATCTTTGGTTA
 CTACAGCTGCTTTCATAAAAAGCCCTCCCTAGTATTAAAGTGTAATCTAT
 TACTTCCACCCTTATTTTACACTCTTTTGAAATCGTTTCCAAAAGTCCTT
 35 TCATTTTACTGAATTTTGTAAATTGTATATAGAGCGACATAATATCGT
 CACTTTTATAGCATATTTTATGATTTATTCATATATTTTATTTTAACTA
 TATATATATATCATGTTCAGTGAAAAGTCATCGCTCTTAATCACACCTTATA
 ATTTTGGGGATGAAGGTGTTTATTTTAAAAATCTTAATTGGAGGGGATAAA
 AGTGAAGACGCAAGTTTTGTACGGAAAATCTGAAATTGAAATCAATGTGCC
 40 TGATGATAGCACAATCATAGAGCCTCAAAACATCGACGCTATTCAAGATTA
 TGAATCAACAATTAATAATGTATTAAGAAACCCACAAATTCTAAACCTTT
 AAAAGAAATGGTAAACAGTAATGATATTGTTTCTATTGTTATTAGTGATAT
 TACACGTCCAACGCCCAACCATATTCTTGTACCTTTACTAATTGAGGAATT
 AAATCATGTTCTCGTGAGAATTTCTGTAATTATTAATGGTACAGGGACTCA
 45 TCGAGATCAAAACGCGAGATGAATTGATTCAAATGTTAGGTGAAGATATTGT
 AAATTCAGTAAAAATCGTTCACAATCATTTGCTCAGAAAAAGAAAGCTAGC
 TAAAGTGGGACACAGTCAATATGGATGTGATGTTTATTTAAACAAAGCATA
 TGTAAGATCCGATTTTAAATTTGTAACAGGTTTATTGAACCACACTTTTTT
 CGCCGGATTTTACAGGTGGACCTAAAGGGATAATGCCTGGAATTGCAGGTTT
 50 AGAAACAATTCAAACATTTTATAATGCAAAAAATGATTGGCGATCCGAGATC
 AACGTGGGGAAATTTAGAAAGACAATCCAGTTCAAGATATGGCACGGGAAGT
 TAACCGTATGTGTAAACCTGACTTTTTTACTTAAATGTTGCATTGAATAAAAG
 TAAAGAAATTAATGCAGCATTTGCTGGTGAATCTTAGATACACACAAAGA
 AGGATGCGCATATGTAAAGATCATGCAATGTTTAAATGTGAGCAACGCTT
 55 TGATATTGTTATCGCATCAAATCTGGCTATCCTTTAGATCAAAATTTATA
 TCAAACAGTTAAAGGGATGAGTGCAGCGAGTAAAGTTGTTAAAAAAGACGG
 TCATATTATTATGGTATCTGAGTGTGCAGATGGCTTTCTGATCATGGTAA
 GTTTGCCGAAATTTTCAAAATGGCAGACACACCTCAAGGTATTTTAGAAT
 TATTACAATCCAACCTTTAAGGAAGTTGACCAATGGCAAGTACAAAAACA

AGCAAGTATTCAAACCTTTTGCCAATGTGCATGTTTATTCAGAACTTACTGA
CCAACAACCTTAAAGACTCGATGTTAATCCCAACCTCTAACATTGAACATAC
AATACAAGAAATTAGAACATCGATATGGCCGTAAATTAACCATTGGTGTAT
GCCACAAGGTCCCTTTAACAATACCGTACGTAGAAGATAAAGAATAATAAAT
5 GGAGGATGGAACTATGACTGAAATTAATACTAAAGAAATGAAATTAGACG
CACTATTGAAAGACATGCAGAGTGTAGTAATTGCCCTCTCAGGTGGAGTAG
ATAGTAGCTTGTACTGAAAAAGCGATTGATATTTTAGGTGTAACTATG
TTAAACCTGTGTAGTAAATCAGAATTATTTAGAAATGAAGAGTTTGAAC
TAGCGCTTAAACTTGGACAAAGTCTAGGTGTTGAAGTATTAGAACTGAAA
10 TGTCTGAACCTCAAGATGCGAATATCGTTAAAAATACGCCGTGAAAGTTGGT
ACTATAGCAAGCGCTTGATGTATAGTCAACTTGAGAATATTAAGAATAAAC
TAGGATTTAATTATGTGCTAGATGGTATGATTATGGATGACTTAGATGATT
TTCGTCCCGGATTAAAAAGCAAGAGACGACTTTGGTGTTCGTAGCGTTTAC
AAGAAGCAAACTATCTTTAGAGCACAGTGGCGATGATATC
15

Sequence 3470

step.1003f04.cons.ok

AGGTGAAGATGTATTCTTAGGATTGACTGTGCTTCTTCTGAATTCTATGA
20 AAATGGTGTATTATGATTACACTAAATTCGAAGGTGAACACGGTGCTAAACG
TAGTGCAGCAGAGCAAGTTGACTACTTAGAAGAATTAATTGGTAAATATCC
AATCATCACTATTGAAGATGGTATGGATGAAAACGATTGGGAAGGTTGAA
ACAATTAACGTATCGGTGATAAAGTTCAATTAGTTGGTGTGATGTTT
ATTCGTAACTAACACTGAAATTTTATCTAAAGGTATCGAACAAGGTATTGG
25 TAACCTCAATCTTAATCAAAGTAAACCAAATCGGTACATTAACTGAAACATT
CGATGCTATTGAAATGGCTCAAAAAGCTGGATATACGCGGTTGTATCTCA
CCGTTCTGGTGAAACTGAAGATACTACAATTGCTGATATCGCAGTTGCTAC
AAATGCAGGCCAAATTAACAGGTTTCATTATCTAGAACTGACCGTATTGC
TAAATACAATCAATTATTACGTATTGAAGATGAATTATACGAAACAGCTAA
30 ATTTGAAGGAATTAATCTTTCTACAATTTAGATAAATAATATCTCTTAA
TTTTTATAAGAGTTCACGCTTATTTAATATCTTAATTGAATTGTCCGT
TGCAGAGTTGAAGCTGCAACGGGCTTTTTTAGTTAAGTACATTTTAAATTA
TTAATGTAACTGATATAATGATGTGAGTAATTTTGTAAGGAGTACAAT
TATGACAGACTCAAATGCTAAAGAAATAAGAACTGGACGTTTAATTGCGAT
35 AAGTTCATTAGTGTGTTTGTATTTTACTTATCATACACCACTTTATTGTATT
AGATGAATCAACAGCTAAATCAATTTTATCTTTAGCTGGTCAAAAAACATC
AGATACAGCAGTGAAAAACATTTTAAATAGTGACCGATACACTGGAATTAT
GTATATTTTAGCTTACTTAGCAGGTACTGTTGCTTTCTGGAATCGCCATCC
ATATTTATGGTGGTTTATGTTTGCCGTATATATTTCTAATGCACTATTTAC
40 ACTCGTAAATCTTACTTATTTATTTCAAGGTATTTTAGATGTAAAAATGT
ACTTGCAGTTTACCAATTTTAAATGTAGTGATTGGATCTATAATTTAGC
AATTTATATGCTAGTTGTTTCTATTACACGTAAAAGTACTTTCAATAGATA
GAAGTTTGATTACATCAATAAAGATGTGTTATAATTTCTTTTCGTATATAAT
GAACGGAGGACAATTTTATGCATACATTAATCATCGTTTATTAAATTATA
45 GATTGTATTGCATTAGTGACTGTTGTATTACTCCAAGAAGGTAAAAGTAAT
GGACTTTCAGGTGCTATTAGTGGTGGCGCTGAACAATTGTTTGGTAAACAA
AAACAACGTGGCGTCGATTTATCTTGCATAGATTAAACAATAATTTTAGCT
ATTTTATTTCTTGTACTTATGTTTGTATTAGTTACTTAGGTATGTAGAGA
ATATTATATTTGTTTAAATGAAGTAAATGAATCTTAACCAAGACTGATAAA
50 TGATAGGGCAAATTAATCACATTTGCTATTAAGGTGATGCTGCCAACCTAC
TGATTTAGTGTATGATGGTGTGTTTGGAGGTGCTCCAGTGGCTTCTGTTCT
ATCAGCTGTCCCTCCTGTTTACGCTACTGACGGGGTGGTGGTAAACGGCAA
AGCACCGCCGGACATCAGCGCTATCTCTGCTCTCACTGCCGTAAAACATGG
CAACTGCAGTTCACCTACACCGCTTCTCAACCCGGTACGCACAGAAAATC
55 ATTGATATGGCCATGAATGGCGTTGGATGCCGGGCAACAGCCCGCATTATG
GGCGTTGGCCTCAACACGATTTTACGTCACTTAAAAAACTCAGGCCGCGAGT
CGGTAACCTCGCGCATACAGCCGGGCGAGTGACGTATCGTCTGCGCGGAAA
TGGACGAACAGTGGGGCTATGTCGGGGCTAAATCGCGCCAGCGCTGGCTGT
TTTACGCGTATGACAGTCTCCGGAAGACGGTTGTTGCGCACGTATTCGGTG

AACGCACTATGGCGACGCTGGGGCGTCTTATGAGCCTGCTGTCACCCCTTG
 ACGTGGTGATATGGATGACGGATGGCTGGCCGCTGTATGAATCCCGCCTGA
 AGGGAAAGCTGCACGTAATCAGCAAGCGATATACGCAGCGAATTGAGCGGC
 ATAACCTGAATCTGAGGCAGCACCTGGCACGGCTGGGACGGAAGTCGCTGT
 5 CGTTCTCAAAATCGGTGGAGCTGCATGACAAAGTCATCGGGCATTATCTGA
 ACATAAAACACTATCAATAAGTTGGAGTCATTACCTGCTATTAAGTATTAT
 CAGCCTGAGTCCGACGATGTACTTGTCTGGACTTTTTTATTTTATAATTGAG
 ACAATCATGTTTTTAAATCATATGTTAAAGGGAATAATGATTTATCAAGTC
 AATTTGATAATAATGAAGAGATTTTTTAAATATATGAAATAAATTGAAATAC
 10 TTAACCTATTTAGGATAAATTGAGTACTTACTATAAGGGAGAGGGATTTTA
 ATCATGCAAATTAACACCAAACCATCTTTTTTGAAGAAGGGAACCGT
 GCAGTGTACTTCTTCACGGCTTTACAGGTAACCTGCTGATGTAAGACAA
 CTTGGGCGTTATCTTCAAAAAAAGGCTATACATCTTATGCTCCACAAATAT
 GAAGGACATGCAGCGCCCCCAGAAGAAATATTAATACTAGCCCTTTTGT
 15 TGGTTTTAAAGATGTTTTAGATGGTTATGATTATTTAGTAGATCAAGGTTAC
 GAAGAAATAGCACTAGCTGGTTTTATCATTAGTGCGCCCTTCGCATTAAAA
 CTAAGTTTAAATCGTGATGTGAAGGGGATTATAACTATGTGTGCACCTATG
 GAGAATAAAACAGAAGGTTTCGATTTATGAAGGCTTTCTTGAATATGCACGT
 AACTTTAAAAATATGAAGGCAAAGATCAACAAACGATTGATCAAGAAATG
 20 GAACAATTTCACTCAACTGAAACCCTGAGAGAACTGAGTGACACTCTAAAT
 GGAGTTAAAGAACATGTCGATGAAGTAATTGATCCAATACTTGTCTGTACAA
 GCAGAACAAGATACAATGATTGATCCTCAATCAGCAAATTATATATATAAT
 CATGTCGATTTCTGATGAAAAAGAAATCAAATGGTATCAACATTCAGGTCAT
 GTGATTACCATTTGATAAAGAAAAAGAGAAAGTCTTTGAAGATGTATATCAA
 25 TTTTTAGAAATCATTTGGAATGGACAGAGTAAAAAATAAAGATGTAGGAGAGA
 AAGGAGGGGCATAATGAATTTAAAGCAATCCATCGAAGAAATGATAAAACA
 ACCTGACTATGAACCCATGTCTAGTATCTGACTTTCAAGATGCGTTAGGTTT
 AAACAGTGCCGACTCATTTAGAGATTTAATTTAAATACTCGTTGAATTAGA
 ACAGTCTGGTTTTAATTGAACGTACAAGAACAGACAGATATCAACGTAAACA
 30 ATCCAACAAAACAAATTCAAAACTAATCAAAGGAACGTTAAGTCAAAATAA
 AAAAGGCTTTGCTTTCTTAAGACCTGAAGATGACGAGATGGATGATATTTT
 TATTCACCAACTAAAATCAATAGAGCATTAGATGGAGATACTGTCTCGT
 GGAAATTCAAAATCTCGTGGAGAACATAAAGGTAAATTTGAAGGTGAAGT
 AAAATCTATTGAAAAGCATTCAGTTACACAAGTTGTTGGAACGTATAGCGA
 35 AGCAAGCATTTTGGTTTCGTATTACCGGATGACAAACGTATTATGCAAGA
 TATCTTTATACCTAAAGGACAAAATTTAGGTGCTGTAGATGGTCATAAAGT
 ATTAGTACAAATTACGAAGTATGCCGATAGTACTGACAAT

40 Sequence 3471

step.1003f07.cons.ok

TAAGCAATAAAGTTCTTCAATTTTATGTAAAATATGGAGGAGAATTCGAAT
 TAAAACAAGGATTTAGTCCAGCGTTCACCGTAGAAAATAAAGATGCTATTG
 ATATAGGTTTTGAACAAACATTCTATGAAATAAATGTCGTTATTGCGGAAA
 45 AAGATTTATGCTATTTTCAAGATGAAAACTAACTGTAGATGCAATCGATC
 ATGAAGATGAGATCATCTACAAACGAAATTAATATAATTATACATATAAAA
 AAAGAAGCGATAAAAACTACTTCGCTTCTTTTTTATAAATGATTAATGCCAA
 ACGTAACCTCAGCATCTTTACCTTCTTTATTTAACTGGTTAACTTTACTAT
 AAGTCTCATGCCACTCAGGAAAATATCTATCCAAGCGTATGGGTTTAAATG
 50 TATCTGCTTTAATACAGATAAGTTCTGTTGAACCAGTAGTAGCTAACTCTC
 CATTTTCAATTATAAACCTCATAACAATAAGTTGAACGTAATCTAGAATATT
 TTTCCACCCATGTTTTTACTGTCACCTTTTCAGGATAGTAAATTGATTTTT
 TATATTGCACTTTTAAATCTACAACTGGTGAAATAACACCTTGTTCCTCCA
 TAGAGGCATAACTGAAGCCAAGCTTCTTATATAGTCTGTTCTCGCAACTT
 55 CAAACCATGTTGCGTAGTTACCGTGATAGATAACCCCCATTTTATCGGTTT
 CTTGATATCTTGCTTCAATTTCAAGTCAAACGTATATCATTGGAATAATCC
 TTCTTTCAACTTTTAACTTATCTTTCAATTAAGGATAACATAAAGCAGTCA
 TGCCAACACAATAGTTAGCATGACTGCTGTACTCAATAGTGAGTCAAAAT
 CATATTAAATTAAGTGGCAAGTTATTTCTTAGTACCATTGTAAAATACC

ACCATGACGATAATAATCTAATTCTACTAGTGAATCAAAACGTACAATAGC
TTTGAAATTAATGATTTACCATTTTCTTTTTTTCAGTCACATTTACAAG
ATCATGTGGCTGTACATCTTCATTAATATCTACAGATATTTCTTCTTTTCC
ATCAAGACCCAGTGCTTCTGCAGATTCTCCTTGTGGAATTGAAGCGGTAG
5 TACACCCATCATAACTAAGTTAGAGCGATGAATACGTTTCATAGCTTTGTGC
AATGACAGTTTAACTCCTAATAAATTGGTACCTTTTGCAGCCCAGTCACG
AGAAGATCCCATTCCATAGTCATTACCAGCTAAGACAACCTAAGCCAGTTCC
ATCTTCTTTATATTTTCATTGCTGCATCATATATAGGCATTATTTCTCCGGT
AGGCCAATATGTTGTAAATCCGCCTTCAGTACCTGGAGCAAGTTGGTTTTT
10 GATACGAATATTGGCAAATGTACCACGTACCATAACTTCGTGGTTACCGCG
ACGGGAACCATAGAGTTAAAGTTGCGAATTGCAACATCATGATCTAATAA
GTATTTTCCCTGCTGGTGTATCTTTACCGATAGCACCTGCTGGAGAAATATG
GTCTGTTGTAACAGAATCACCAAATTTACCCATAACTCTCAAACTTTAAAG
TGGTTCAATTTTACCCGGCTCTTAGATAATCCTTGGAATAATGTTGGATT
15 TTGAATATATGTTGAATTAGGATCGAAATCATATAATGGTTTCATCGGTTAC
ATCTATTTTCATTCCACATTTTCATTGTTATGATATACATTTTATATTTCTTC
TAAGAATAGTTTCAGGTGTAACAACCTTTATCTACAGTATCAGAACTTCTTG
TATTGAAGGCCAAATAGTCTTTGTAATATACATCTTTACCATCTTTACCTTT
ACCAATTGGTTTCATTTTGTAAATCAATATCTACTGTACCAGCAAGCGCATA
20 AGCTACAACAAGTTGTGGTGAGGCTAAATAGTTTGTCTTCACTAATGGATG
GATTCGCCCTCAAAATTACGATTACCTGATAAACTGAAGTTACTAATAA
ATCTTCATCCGCAACTGCCTTTTCAATTTTCAGGTAATAGTGGCCCTGAGTT
ACCAATACATGTAGTACAACCATAACCAACAAGATTGAAACCTAAATCATC
TAAATACTGTTGTAATCCAGAATCTCTTAAATATCCTGTAACAACCTTTGA
25 ACCTGGAGCAAGTGACGCTTTAACAACTCTGGTACTTTCAATCCTTTTTC
TACAGCTTTTTFAGCAACTAAACCAGCACCTAACATAACATATGGATTAGA
AGTGTTAGTACATGAGGTAATAGCAGCAATTGCTATATCTCCTGTTTTCAT
TGTTGCTTTTATAGTCCATCATTTAAATTAATTTCTGCTTTTTTATCAAATTC
ACTTTGATCAAGTCCATGTCTTGATTACCAGCAGGAGCAGTTACTGATTT
30 TTCAAATCTTTTTTTCATATCACTTAAGAAATTAAGTCTTGTGGACGCTT
TGGACCAGAAAGTGATGCCTCTACTGTAGATAAGTCTAAATCAATAACATC
TGTATATTCAGGATCCTCTTTTCTACATCAAAGAACATATGGTTTGTG
TAAATATCTTTTAAACCAATTCATATGTTCTTCGTCACGGCTGTAAAGTT
CATATATTTCAATGATTCTTCATCTACTGGGAAGAAACCACACGTTGCACC
35 ATATTCCTGGAGCCATGTTAGCAATTGTAGCTCTGTCTGCTAATGGTAAATG
TTGAACACCTGGACCGAAGAAATTCACAAATTTTCCAACCTACACCTTTTTT
ACGTAATCTTCAGTCACACGTAAAGCTAAATCCGTAGCTGTTGAGCCTTG
TGGTAAAGAGTGTAGTTAAACGCACTCCGATAACTTCAGGAATTGGGAAATA
TGATGGTTGTCTAACATACCTGCTTCGGCTTCGATACCACCAACGCCCCA
40 ACCTAGAACACCAATACCATTAAATCATTGTAGTATGTGAATCAGTACCTAC
TAAAGTGTACGAAATGCTGTTTTCACCATCAACATCTCTTACATGTAC
TACATTTGCTAAATACTCTAAGTTTACTTGATGGACAATACCTGTAGCAGG
AGGTACTGCATTATAGTTATCAAAAGCTTTTGTGCCCAGTTTAAAAATTG
ATAACGTTTCATAGTTACGTTCAAATTCATTTTCATATTACGTTCTAATGC
45 TTCTGGATTAGCGTAACATCAACTTGAACCTGAATGGTTCGATAACTAAATC
CACAGGTACTTCTGGGTTGATTTTATTAATATCTCCACCAACATCATTCAT
AGCTTTACGTAAAGAAGCCAAATCTACTACTGCTGGCACACCTGTAAAGTC
TTGTAAAATAACTCTAGAAGGTTTGAATGGAACCTCACCTTCGTTACCTGC
ATTTCCGAATTTACTTAATGCTTTGATATGATCATCTGTTATAACAAAATC
50 ATCCTCTGTGCGTAACACAGATTCTAACAATACGCGAATTGAGTATGGTAA
TTTA

Sequence 3472

55 step.1003f08.cons.ok

AGCCACTGACCGTATTTACGGCCAGTGGCTTATTAAATATAATATAAAAAA
CCGAGCGTATAAACGCTCGGATATAATGATTTATTTAATTGTAAATCAAC
TTCATTTATTTCTTTTAAATTTTCAGCTGTATCATAGCAATTTGTAGCGAT
ATTTGCATATTTCTCTGCTAAACGATAAGCATTAATGTATAAAATTAACT

TTCTTTAGCGATATCTTCAGCGTCATCAGATTTGGCTGGATTTGAAGTTAC
 AACTAATTCAGCAAATTTTTCAGGATCAATATTAATAGACATGAATGAATA
 AACAACTCCTATATTTTATTAGTTTCATATTAACACAATACTTAAATACA
 AAAAAAGTAAAGTCTTTTAATTTTCAGAATAATTAATTTGTTAATGAGACTT
 5 TTATTATAGTACGGTATCCATTAACTTAATTTATGAGAGGATGAACAGAG
 AATGAGAATATTGTTTATAGGTGACATCGTTGGTAAAGTGGGCAGGAAAAAT
 GATTACTACTTATTTACCTAAAAATTAAACAACTTATCACCCAACAGTTTC
 TATAGTAAACGCTGAAAAATGCCGCACACGGTAAAGGATTAACAGAAAAAAT
 TTACAAACAACTTTGTAGAGAAGGCGTGGATTTTCATGACTATGGGTAATCA
 10 TACATATGGTCAAAGAGAAATTTACGATTTTATTGATGATGCTCATCGAAT
 GGTGAGACCTGCAAATTTTCTGTATGAAGCTCCAGGAACAGGTATGAGAAT
 AATAAAAAATTAACGATATTAATTTGGCTATTATTAATTTACAAGGCCGTTTC
 ATTTATGCAAGACATTGATGATCCATTTAAAAAGGCTGACCAGCTAATCGA
 AGAAGCTCAAAAATCTACACCATATATATTTGTAGATTTTCATGCTGAAAC
 15 TACATCTGAAAAAATGCTATGGGTTGGTATTTAGATGGTAGAGTGAGCGC
 TGTGTTGGTACTCACACACATATTCAACTTCTGATGATCGTATATTACC
 TCATGGCACAGGATATATCACAGATGTCGGGATGACAGGTTATTACGATGG
 TATTTTAGGTATCAATAGAGATGAAGTTATTCAACGTTTATTACTAGTTT
 GCCACAAAGGCATGTTGTTCCAGATGATGGGCGAGGCGTATTATCAGGAGT
 20 TATCATAGATTTAGATAAAGAAGGTAAAACGACTCAAATAAAAAGACTGTT
 AATAAATGAGGACCATCCTTTCCAAATTTAAAGGGATATAGTCCATTAGTA
 TGAATTTAACTTGTACCACTGTTGAAATAGTAAACGGTGGTTTATTTTG
 TTATCATAGATGAAATTAATACCACAGGAGGCATGTGATATGAAATCAC
 AAATATCATGGAAGTGGGCGGTACGAAGGCGAAGGTATTGAATCTACCG
 25 GTGAAATCTTTGCTACTGCGATGAATAGAAAAGGTTATTTTGTATGGAT
 ATAGACACTTTTCTAGTCGTATAAAAGGTGGCCATACTAATAATAAGATAA
 GAGTTTCAAATCGCCTGTGCATGCGATTAGTGATGATTTGGATATACTCA
 TTGCTTTTGACCAGGAAACGATTGAATTAATCATCATGAAATGAGAGAAG
 ATAGTATTATAATTTGCGGATGCTAAAGCAAACCCCAAAGCCAGAGAAT
 30 ATGTGGCTCAATTAATTGAGTTACCATTCCTAGCACGGCAAAGGAACTTG
 GAACAGCATTAATGAAGAATATGGTGGCAATTGGTGCGACATCTGCACTGA
 TGGATTTAAATACATCAACTTTTGAACTTTAATCGATAACATGTTTCAA
 AAAAAGGTAAATAAGTCGTTGATATGAATATACAAGCCCTTAATATGGGTT
 ATGATTTAATGAAGCAACAAGTTACCAACGTTAATGGAGACTTTACATTAG
 35 AGAATGGTAGCGGTATCCTCATTTATATATGATAGGTAATGACGCAATCG
 GATTAGGAGTATAGCAGCTGGATCAAGATTTATGTCCGCTTATCCAATTA
 CGCCAGCTTCTGAAATTATGGAATACATGATTGCCAATCTACCTAAAGTTG
 ATGGTACTGTTGTTCAAACGAAGATGAAATAGCAGCAGCAACGATGGCGA
 TTGGAGCTAACTATGCTGGCGTACGAGGCTTTACAGCGAGTGCGGGTCCAG
 40 GTCTTTCTTTAATGATGGAATCTATTGGATTGCTGGTATGACTGAAACGC
 CATTAGTCATTATTAATACTCAAAGAGGTGGTCTTCTACTGGCTTACCAA
 CAAAGCAAGAACAATCAGATTTAATGCAAATGATTTATGGTACCCATGGTG
 ATATTTCCGAAAATTGTCTGTTGCTCCTACAGATGCTGAAGATGCGTTTATC
 TTACTATGGAAGCATTTAATTTAGCTGAAGAATACCAATGTCCAGTCATTC
 45 TGTTAAGTGATTTACAATTATCATTAGGAAAACAACTGTTAAACACTCG
 ATTATAATAAAATCGATATTCGTCGTGGAGAAATAATACAGTCAGATATCG
 AGAGAGCTGAAGATGATAAAGCATACTTTAAAGATATGCATTAACAGCTA
 GTGGCGTATCACCACGACCAATACCAGGTGTTAAAGGTGGTATACATCATG
 50 TAACA

Sequence 3473

step.1003f10.cons.ok

ATGGTACGTATGAAATGGCAAATACCCTAGCTGAAAACAATAATCCCTATG
 55 CATATGTTGTCTAGTGCATTAAAGAAAACGTTAAAAGCGTGTATCGTGCCCTT
 TGAAAGAGGTGGAATCATGCTTAATCAATTTGTATGGGTTATTTATCCA
 TATTTATGTTTAGCAATATTTGTTATTGGACATATCGCAAGATATAAATAT
 GATCAATTTTCATGGACAGCAAAGTCGAGTGAAATGATTGAAAAGAAAACGA
 TTAATAATGGGGAAGTTTACTTTTTCATTTAGGTATTATTCCGGTATTCCTT

GGTCATGTTGTGGGATTATTGATTCCCGCTAATTGGTTAGAGGCGATAGGA
GTAAATAATCATATCTATCATATAGGTGCAGTTTATATAGGTAGTGTTTT
GGTATAATAACATTAATAGGAATGTTGTTGCTAACTTTAAGACGACTATCC
ATCAAAAACGTTAGACGATTAAAGTTCATTTTCAGATATATTTGTGAATATC
5 GTTTTGTGTGATTATTTTAATAATGGGTGTTATTCTACGCTTGTAAACCAAT
GCGATTCAACCTGAATTTGATTATCGTCAAACCATGCGATATGGTTAGA
CATTTATTCATGTTTTCTCCAAATGCTGACTTAATGTTAAACGTGCCCTTGG
TCGTTTAAACTGCACATATTTATTAGGGTTTACAGTGTGTCGTGTTGGCCA
TTTACTCGTTTAGTACATGTTTGGAGTGTACCACTGTCTTATATGAACAGA
10 AGATATATGTTTATCGCAAAAACAAAATTTAATTTATTATGAGGTGAAAG
TGTGGAACCAGAATCTTTAAACATAATCAACATTTACAAATGAATTTAGA
TAAGTTGAGAGCACAAGAGGGTTATGATTTTGGTGGTATCGCTTTATATGA
TTATCATCACACTTCATCACCAATTAAATGGCAATATGTTTCAGGTAAACAC
AAATGATAGATATAAACTTATCATTTTGAGAAAGGGTAGAGGGCTTGCTGG
15 AATGGTGATGAAAACCGGTAAAGCGTATGGTTATTGCTGATGTAGATACAGC
TTTATCTCCAGAAGAGAAAAGTTAAATTTCCAATCATCTTAGTGAGTCATT
GACAGCTGTAGTTGTCAGTCCCTTTATGGTTAGAAAATTCATGTATGGCGT
TTTATTATTAGGTCAAAGAAATCATCAGCCGTTACCTCAGTCATTGGACCA
ACTTAATATTGAAAAACAAATCGGTATTTTACAGAAATAAATAGGTGGT
20 AAATATGTTGGAGCAAACCTGATTAAAGTTTAGAGCAATTACTTAAGAATTA
TTATGAAACCACGAACGAGAAAAATTGATTTTGTAAATAGACAAGGCAAAAT
TATTGCTATGAATGACGCAGCAAAAGATATTTAACTGAGGAAGATAATTA
TAATGCTATGACAAATGCGATTGTCATCGATGCGAAGGATACCTAATGA
ATATGATGTACAATCGTGTAAAGATTGTTTTTATAGACAACGCAATTACA
25 ACATTTCCAATTTCCAAGTATTTATGAAGACAAAAGATAATGAAATTAAGCC
TTTTACAGCTATGTATCAAAATATTGATGAACAAAGAGGTATTAGTGCATT
TACCTTACAGAATGTGGCGCCTCAGATTGAAAGGCAAGAAAAATGTATCA
ACAAAAATGTTACATCGTTCAATTCAAGCACAAGAAAATGAACGAAAGCG
TATTTCTAGAGAATTACATGATAGTGAATACAGGATATGCTCAATATAGA
30 TGTGAACTAAGGCTTTTGAAGTATAAGCACAGGATAAGGTGTTAGCTGA
AACATCTCAACGTATAGAAGGCTTATTATCACAGCTTATTGATGATATTAG
AAATATGCTGTGTTGAATTAAGACCTTCTCTCTCGACGATTTAGGCATTGA
AGCAGCTTTTAAATCATATTTTAAACAGTTTGAAGAAAATTATGGTATGCA
TATTAAATATGATTGCAACATTAAGGCGATGCGTTTGTGATAATGAAATTGA
35 AACAGTTGTGTATCGTGTAGTTCAAGAGGGTGTATTTAATGCTCTAAAATA
TGCTGAGGTAAATGAAATTGAGGTAAGTACGCATAGTGATGGCAAGCAGCT
GTGACGAGAGTTGTGGATCGAGGTAAAGGGTTTATGTTAGATCATCACCC
TAAAGGCTCTGGACTTGGATGTACGGAATGAGAGAACGTGCAGAATTAGT
TAACGGTCATGTTAATATAGAGACACATATTAATAGAGGTACTATAATTAC
40 ATTAGATATACCGATTTAAGTTGATGATGTAAGTTGGGGGAATTGAAGTGA
AAATAGTTATAGCGGATGACCATGCAGTTGTTAGGACAGGATTTTCAATGA
TATTAAATTATCAAGAAGATATGGAAGTTGTTGCAACTGCAGCTGACGGG
TTGAAGCTTATCAAAAAGTGTAGAACATCGACCAGATGTTTTAATTTTAG
ATTTGAGCATGCCGCCAGGAGAGTCAGGCTTAATCGCAACCAGTAAAAATTT
45 CTGAAAGTTTTCCTGATACATAAAATTTAATACTTACGATGTTTGATGACG
AAGAATATTTATTTTCATGTGTAAAAAGTGGTGTAAAGGATACATTTTAA
AAAATTCACCTGATGAGCAATTAATATTGGCCGTACGTACAGTATATCAAG
GTGAACTTATGTTGATATGAAATTGACGACGTCCTTAGTCAATGAGTTTG
TCAATCAATCACAAACGGATGAAGTGTATCATCTTCAGATCCATTTAAAA
50 TTTTATCGAAACGAGAGTTAGAAATATTACCTCTTATAGCAAAAAGGCTATG
GCAATAAAGATATTGCAGAAAAGTTGTTTGTATCGGTGAAAACGGTAGAGG
CACATAAAACGCATATTATGACGAAACTAAATTTAAAGAGTAAACCTGAAT
TAGTTGAATATGCCCTTAAAGAAAAATATTAGAAATTTAAATGTATATCA
TGATTGATTTATTAGGGCTTACTTAGTCTATGACTAAGTAGTCCTTTTGTG
55 TACCTTTATATAGGGAAAAATAAAAGTAAATCAGGGTATGTCCTTATGTT
CATATTTTACAAACATATTAATAATGAATTTGTGATGGTACTTTGTTTAT
AATTTTGTTTTATATTATTAATTTTGACCTCTTAT

Sequence 3474

step.1003f11.cons.ok

5 ATATTCAAGTTGGTCAATTTGGAGAATTAAATTTAAATCACTGCTAGAGC
CAACACATATAAATTCAAATGCATGTGGCACTTTGCATAGTAATATAAATC
ATCAAATTCATCGAAAATCCTAGGCTACAAACAAAAGAAGAAATGATTAGTG
CGTTAGATAACTTGCCCAAGATGTTTACCGTGTCAAAGGGTTTGTTCGTT
10 TTTACAGATCAGCAACACGTTTATTTAGTACAGTATGCACAAGGAAATATAG
AATTATCTCCCATTTCAACTTAAAAACGATGTACCATTGTACCTCATTTGTTA
TAGGAAAACATTTAAAAACAAATACAATTTGATTTATAAAAAACACATCTCTC
AATTATCATCAGTTAACTTTATACATAAAACAACTCTAAATGACTTTTG
TTCATTTAGAGGTTTTTGGTACTTATTTCTATTGTTTCGTCTTTTATTAATG
15 TTTGCGCTTTTGAACGTATTCATGCTTACTTCTACTCATTATTTACCTCA
CTTGGATTAAATTCCAAATATCCGCAAAGTAGACGCTTTCTTCTTTATTG
TCGACAAAAGAAAACGATATCTTTATAACCTATAGATGAACCTTTAAGG
TCACCGTCCCTTTTAAATATTAGTTTAGGAGCTCTCTTAATCGGAATATCA
TACCTCTTTCTGAGTTGCTTCACATTGTATCTTCATTACTTAATTGGTAC
TCTGCCGAATAGCTTGGTACGTTTGGGTTATAAGACACATTTCCATCTTG
20 TAGTCTTTTAAATTTTTTAAATTTCCCATATTGAGAGAAAACGAAAGTTT
TTAATTTTCATTTTTTAAATTTTTCATCGGCGATAGGTTTAGTCGGTTCAATT
TTATTGTTTTTAAAGCGAACCGGATATTCTTTATCTTTACTGTGCACTCTC
CCTTTTTTACATCTCTCTAATTATATTGTATATAAAATGACCAGTCGTCTTT
CTAGTATTTCTATTATACATAACAAACCATACCTCTAGATTTTCATCGCTTGA
25 TCTTTTTTCTGAATATTCATTTCTGAATTAATAATCCATGTCCCTTTGTCC
TCTTTTTCAAATTTCTTCATCAGATAGCCTTCTTTATCGTATAAATCTTCT
AAATTTTTTAAATTGGATACATCTTAATGTTTTATTAATTTTTCTTTAATT
TCCGCTTCTTTATTATTTTCTTTATTTCATAAATCCACATCCGCCAATAAAA
AACGTTAATATTAATAATGCTTATGTAAGAAATGCACTTCCCATAAATTTTC
30 ATCTTTTACATCTCTCTATTATTGTATACATTTATTAATATACATTTTA
TTGATATAAATATCAGTTTATAAAAAATTCATTTTATATCATCAAATTAAT
CATACACATTCATTTAATTTATTTATATGTTTTTTTCGGCGAAACAATCA
AAAAACCTCATTAATGTACCGCAATCATACATCTAATGAGGCTATTATCG
TTCTTTTAAATCGCTACTTACGTGTAATACCATCTAATCTTCGACTTTCAA
35 CATATTTATTTATCATCTCATATAGAATTGACCTAATCTTCTCCAGTTTCG
CTTATAATCAGTAGTATCTAAGCCGACTAAATCTAGATGATCCCAATCATA
TTGTACGGGTCTCACTTGCCAAATACCTTTGTCAGTTGCTAGGTTTCATCAT
ACCTACCTTCTTAAATGCTTCATCACTTGGATGTAGTGAAGAAGACACAGG
TACTATGCCATCATTTACTCTGACATTTTTATTTATCATCTCCACCTATCAC
40 ACGACTTGTAAATCGAATAGTGGGAATTGTCTAATATTTCGGCACTTCATT
GCCTAATGGTCCAGTATGTGTGTCAGCACCTGTATATGATGTATAGACGAT
ATTAGGATTCATGTGCTCATTGGTTTAACTTTTCTGCTCCAACAGTTGT
TAAATCATTTACAGCTGATCTTTCAGTCTCCCAAACCTTTACTATTGCTAT
ACGTTTTGCATATTCAGCATATGATTTCATTAGGTTTCTGTTTGAAGCCCCA
45 TTGAGAAAAACCTAGTTCTAAATCGAGCGCTTTAGTTCCACCAATTTTTCC
AATTCTATTAATTGTATCTTTGATAAATTTAGTTCGACCCTAGTTTATCTGC
AGCAGGTGTGCCATTATGAGGTGTTCCCTAATGTAGTAATCGTAGACACCAT
GTTATCTTGTCCACCTTTAAACAAATCAGATACCGTACCACCATATTGACG
TTGATAATCTATTTCTTCTTGATTTCCATTTCTTAAAAAATGTTCCATCAA
50 GCGTATCGTTTGGCCACCATACTATGTCCAACAAGATGTATCTTTTACC
TGGTTCCCAATCAGGCATGATGCCTTCATATGTTCTGCCATAACGCTTGTG
ACCATATTTTGCAGCATGTGCTGCACCATAATCTACTCTTCCACCTTTAAT
ATAATAATACAGTTCAACAGCACGGTCATAATTGCTGCTAAATGCTCCTAC
ATTGGCTTCGTGAACCTCGGTAACCTAATTTGTAAGTTCTTGTTCACGTT
55 ATATTTAGTACCACCCCAATAATTTGGGTACATGCTGAATGAATCTTCACC
GACTAAACCTACAAATCCATGTACAAACACGACTGGATATTGATTTTATA
TTGAGCTTGGCAGTAAGTTGATTGATATGTTTAGTTTGTGTTTGAATTCAC
AACGCGTGTGAAGTTAATGATGTTTGCTTTTTTTGATTTGTTTTAACTGT
TTTATCAGTTTTTGTATTACCTATTTTGTGTACTTTAGAAGATGGTTGTTT

5 TACTGTTTATCTCTATCTCCTATAGACTGTGATTGTTTAGGTTGTGTTCC
TTTTGAAGATGGGGCTTGAGTTTGATCAGATGGGTGTGCTTTTTATCTGA
TGAATCTTGATCATTATTTTGGTTTTGAGTATTAGGTTGATGTTCTGTTTG
TGAGTGTGATTATCTTTAAACTGTTGTTTCTTTATCATCTTGATTTTG
10 GCCTTTTGGATATGTATCCTGACGTGAATTTGGGGTAACATGATTGTCTTT
TGATTGTTGCCCCCTTTAAAGTTTGTTCTTGGTCATTTAATGGTGACGTTTG
TGTTGAATCTTGTTGAGTTTGATTATTCTTTTTTGTACATTTTCATCAGA
TTGGTTCCTTGATTAGTCTTTGATTGGCTAGCTGTGTCACCCTCTTGATT
CAATTCGTATGATAAATGAGTTGTATCTTCTTTGTGTTCAACTTGTCTGT
15 TTGTTGTGACTGTTTCGTTAAATCTGATGATGATTGTGTCAATTCAGCAGC
TTGAGC

Sequence 3475

15 step.1003f12.cons.ok

AACGAAATACAAGAAAATGTGACTGAAAAGACGAAAGGTCATAAATTAAAA
AAAAGTGCAGCTAAAACAACAGCCCTAGTTGGTGGAGCATTACATTTAAT
ATGTTGAATAATCATCAAGCATTAGCTGCCCTCAGAAACACCAATCACCTCT
GAAATTTCAATCCAATAGCGAGACAGTAGCCAATCAAAATTCAACTACGATT
20 AAAAACTCACAAAAAGAAACAGTCAATTCTACAAGTTTGAATCTAACCAT
TCTAATAGTACAAATAAGCAAAATGTCCTCAGAAGTTACAAATACAGCTCAA
TCAAGTGAAAAAGCTGGAATTAGTCAACAAAAGTAGTGAAACATCAAATCAA
TCATCTAAGTTAAATACATATGCCTCCACAGACCATGTAGAGAGTACAAC
ACTAACAATGATAATACTGCACAACAAGATCAAAATAAGTCTTCGAATGTA
25 ACCTCTAAGTCAACACAATCAAACACGTCATCCTCAGAAAAGAGCATTAGC
TCCAATTTAACCAGTCAATCGAAACAAAAGCAACCGATTCTTAGCGACC
AGTGAAGTACGTACTAGTACAAATCAAATATCTAATCTGACATCAACATCT
ACTTCAAAATCAATCGAGTCTACTTCTTTTGCAAATTTAAGAACATTTAGT
AGATTTACCGTTTTTGAATACGATGGCAGCACCGACAACAACGTCACGACA
30 ACAACTTCAAGTCTGACATCTAATTCTGTTGTGGTGAACAAAGATAACTTT
AATGAACATATGAATCTATCTGGATCTGCGACGTATGATCCAAAAACAGGT
ATTGCTACCTTAACGCCAGATGCATATAGTCAAAAGGGTGCCATATCTTTA
AACACTCGATTAGATTCAAACCGTAGCTTCCGTTTTACAGGTAAGTTAAC
CTTGGTAATAGATATGAAGGTATTCTCCTGATGGTGTAGTAGGTGGAGAT
35 GGCATTGGCTTTGCAATTTTACCAGGTCCCTTAGGACAGATAGGTAAAGAG
GGGGCTGCCGTTTGAATAGGTGGTTTGAATAATGCCTTTGGATTAAATTG
GATACGTATCATAACACATCACCTCCTAAATCTGATGCTAAAGCAAAAGCA
GACCCACGTAATGTTGGTGGTGGTGGCGCTTTTGGTGCCTTCGTAAGTACA
GATAGAAATGGTATGGCTACCACTGAGGCATCATCTGCGGCTAAATTAAT
40 GTACAACCTACTGACAATTCAATTCGAAGATTTTGTCAATTGACTATAATGGT
GATACAAAAGTGATGACAGTGACGTACGCTGGACAACTTTTACGAGAAAT
CTTACAGATTGGATTAAAAACAGTGGTGGTACGACGTTTTCTCTATCTATG
ACTGCCCTCAACTGGTGGCGCAAAAACTTACAACAAGTTCAATTTGGAACA
TTCGAGTATACAGAATCAGCTGTTGCTAAAGTACGCTATGTAGATGCAAAT
45 ACTGGTAAGGATATTATTCACCTAAAACCATTCAGGTGAAGTTGATGCG
ACTGTGAATATAGATAAACAATTAAACAACTTGAAAAATTCAGGTTACAGT
TATGTTAGTACAGACGCTTTACAAAACCTCCAATFATTCAGAAACATCAGGT
ACACCTACACTTAAATTAACTAACTCAAGCCAAACGGTGATTTATAAATTC
AAAGATGTTCAAGGTCCCTCAAATTAGTGTGATAGTCAAACTAGAGAAGTT
50 GGAAAGACCATTAAATCCAATTACAATTACTACAACCTGACAATAGTAAAGAC
GTATTAACATAAAGTGTGACAGGTCTACCTTCAGGGTTATCTTTTGATCAA
ACGACAAATACAATTACTGGCAGCCAAGTGAAGTAGGAACATAAAGTGTG
ACAGTTAATACTACTGATGCTACTGGGAACGTAACATCTAAGCAATTTACA
ATAACGATTCAAGATACAATCAGCCCTGTTGTAAATGTGACGCCAAGTCAA
55 GCATCAGAGGTCAACGACAACAGCTATAAAGAACCATAAAATATTACTTAT
TCCACCTAGTACAGCGACAAATAGGACCTCCATGAGCAACTTTATCACCAGC
ACCTCCAACCGCGCTATTACTGAAGCAATCATCGCGCCAATCATATTGGC
TGGAATGATTCTTAGAGGATCTTGTGCAGCAAAAGGGATTGCCCCCTCAGT
AATACCAAATAATCCCATCGTAAATGATGCCCTTACCATTCTTGTCTGCTGC

TTTATTGAATTGGTGTTTACGAACAAATGTAGCTAAACCTAAACCAATCGG
 TGGTGTAATACTGCTACAGCAACCATTCCCATCACAGTGTAGTTGCCCTC
 AGCAATTAATGCAGAACCGAATAAGAATGCTACTTTGTTTACTGGACCCTC
 CATATCAAAAGCAATCATCGCGCCAATAATAAGAGCAAGAATAATGATGTT
 5 AGCACCTTGCAATTCCTTTTAACCATGATGTTAATGCACCAAAATATATTTGA
 AATTGGTGGCCTATTACAAATATAAAAAATGAGACCAACTATTAAAGAAGA
 TAGAATAGGTATAATAATAATAGGCATAATAGGAGCCATAGCTTTAGGAAC
 TTTAATCTGTTTAAATCCATTTTGCATATAGCCCGCTAAGAAACCTGCGAC
 10 GATACCACCAAGGAAACCGGCTCCTGCTTCACTTCCATATAAACTACCATC
 AGCAGCAATGGCACCACCAATCATACCTGGAACAAGACCAGGCTTATCAGC
 AATATCACCAGCATATAACCAGCAAGGATGGGAACCATGAATTTAAACGA
 TAAACTACCAATATTTTCAATAGATTTCCAAAATGAATCTTCGGGGATAAC
 TAATCCTTTTGGAGTGGTGTGTCCTCCAAGAGTTAAGGCAATAGCTATGAG
 TAAACCGCCAACGACGATAAATGGAACCATGAAGGAAACACCATTCATTAA
 15 ATGCTGATATACCATTTGAATATTACTTTTATGAGATT

Sequence 3476

step.1003g01.cons.ok

20 CGCCAAAACATGATATTGAATGGACATGTTGATGTAGCTTCTGTAGATGACG
 ATCAATATTGGCAGTATCCACCTTTTAAACTTACCAACAAAGAGAATGGT
 TATACGGTCGTGGCGTTAGCGATATGAAAGGTGGTATGTCTTCAATTATTCT
 ACGTCTTGGAGCAATTACATCAAGAGGGGCAACGTCCAGAAGGTGATATTA
 TTGTTCAATCAGTAGTGGTGAAGAAGTAGGTGAAGCAGGAACATAACGTG
 25 CATGTGAAATAGGACCTAAAGGTGACTTAGCCCTTGCTTAGATACGAGTG
 AGAATCAAGCACTTGGGCAAGGTGGCGTGATTACCGGATGGATTACAGTTA
 AAAGTAAAAATACAATACATGATGGTGCAGTAGTCAAACGATACATGCTG
 GTGGGGGCTTGTGTTGGTGCAAGTGCCATTGAAAAAATGACAAAGGTATTCT
 AATCGCTTAATGAACCTTGAAGGCATTGGGCTGTCTATGAAGAAGAGCCCTG
 30 GAATGCCCTCCAGGTGCGAATACAATTAACCCAGCTGTATAGAAGGTGGAC
 GTCACCTTCGATTTATTGCGAGATGAATGTCGATTATGGATTACTGTTTATT
 ACTTACCGAAGCAAGTTATGAATCTGTAGTTAATGAAATAGAGCAATATT
 TAAATAAGGTGCGAAGCAGATGTATGGCTCAGAGAGAATCCACTTGAAT
 TTGAATGGGGTGGTACATCCATGATTGAGGATAAAGGAGAAATCTTCCCAA
 35 GTTTCCTGTTCCGACACATCATCCAGGTTTTAAGCAATTAGAAGAAGCAC
 ATGAACATATTCATAATAAAAAGCTTGAACATGGTATGAGTACAACCTGTAA
 CTGATGGAGGTGGACAGCACATTTTGGCATTCCCACGATATTATATGGCC
 CAGGTAGTTTAGAAGAGGCACATAGTGTAGATGAGAAAATAAAAGCAAAGG
 AATTAGCTCAATATAGTGATGTTTTATATACATTTTAAAAAGAGTGGTATG
 40 CACACCCACAATCCTATAAATCATCATAGATAAAAAAGAGGTACAAGCAG
 ACCTTTTATACTCACAAAATGTTGAGTTAAAGTGATGTGTTCTGTACCTCT
 TAATTATTATAAAGAATAACAAGTTGTATATTTATAAAAAGATATTATTA
 ATTTCTGTGTTTCAATTCATATAAAATACAATTGTCAAGGTGAGTTTATCTTT
 TAAATGTACTTTTAAATTTTCGTATCCAACCTACCACCACGTCTAGAATGAA
 45 CGGGTTGTCCAGAAGGTTGTACCGATTGTTTTTTTAGTTCTTCTCCATACT
 TAATTGCTTCTTGCATCAAGTAAGATTGAGAGTTTAAAGGAGATGAATTGT
 TTTTCGACATAATGATAAACGCCCTTGTGCATCTTGGTAACGACCTTTTGGT
 TGTCAGATAAATTGTAAATTCATATAGTTAACTAATCGTTGTCCTATTGATT
 TATCTTCAACAGGGAATAATATTTCGACACGTTTAAATCATATTACGTGTCA
 50 TACCTCAGCTGAAGATAAATAAATATGCGCCTCACCATTATTATGGAAGT
 AGTAAACAGTGAATGTTCAAGCAAACGACCCACAATACTAACACCTCTA
 TATTTTCGCTAATACCTGGAATGCCTGGTTTAAAGACAACATATACCACGAA
 TGATGAGTTTATTTTAAACGCCCTGCTTGGGATGCTTCGAAGAGCTTTTCGA
 TAATCGTTTTATCGGTTAAAGAGTTCATTTTCATCATAATTTTACCGTTAC
 55 CATGTTGTAAATGACTACGTATTTCTTTATCGATACGATCAATGAAGACGT
 CTCGAATATCGTATGGTGCTACAATCAATTTATTGTATTCTGGTTTTGTTG
 AGTAACCACTCAAGTAATTAAAGAAGTTAATTGCATCCTCAGCGATATCTT
 TATTTGTCTGTATGATACCCATATCTGTGTATAATTTAGCAGTTTTATCGT
 TATAGTTACCTGTGCCTAAATGAACAAATGACGTAAGTTTATTGTTGATGC